(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
7 September 2001 (07.09.2001)

PCT

(10) International Publication Number
WO 01/64885 A1

(51) International Patent Classification[7]: C12N 15/12, C07K 14/47, 14/705, C12N 15/62, C07K 16/18, 16/28, G01N 33/53, A61K 38/17, C12Q 1/68

(21) International Application Number: PCT/US01/06891

(22) International Filing Date: 2 March 2001 (02.03.2001)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/186,462 2 March 2000 (02.03.2000) US

(63) Related by continuation (CON) or continuation-in-part (CIP) to earlier application:
US 60/186,462 (CON)
Filed on 2 March 2000 (02.03.2000)

(71) Applicant *(for all designated States except US)*: **AMGEN, INC.** [US/US]; One Amgen Center Drive, Thousand Oaks, CA 91320-1799 (US).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: **ZHANG, Ke** [US/US]; 2146 Peak Place, Thousand Oaks, CA 91362 (US). **LINH, Cam** [US/US]; 605 Hazelwood Way, Unit D, Simi Valley, CA 93065 (US). **NAKAYAMA, Naoki** [JP/US]; 325 Briar Bluff Circle, Thousand Oaks, CA 91360 (US).

(74) Agent: **NOONAN, Kevin**; McDonnell Boehnen Hulbert & Berghoff, 300 South Wacker Drive, Suite 3200, Chicago, IL 60606 (US).

(81) Designated States *(national)*: AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZA, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

**Published:**
— *with international search report*
— *before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments*

*For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.*

WO 01/64885 A1

(54) Title: CHORDIN-LIKE-2 MOLECULES AND USES THEREOF

(57) Abstract: The present invention provides Chordin-Like-2 (CHL2) polypeptides and nucleic acid molecules encoding the same. The invention also provides selective binding agents, vectors, host cells, and methods for procuding CHL2 polypeptides. The invention further provides pharmaceutical compositions and methods for the diagnosis, treatment, amelioration, and/or prevention of diseases, disorders, and conditions associated with CHL2 polypeptides.

# CHORDIN-LIKE-2 MOLECULES AND USES THEREOF

## Field of the Invention

The present invention relates to Chordin-Like-2 (CHL2) polypeptides and nucleic acid molecules encoding the same. The invention also relates to selective binding agents, vectors, host cells, and methods for producing CHL2 polypeptides. The invention further relates to pharmaceutical compositions and methods for the diagnosis, treatment, amelioration, and/or prevention of diseases, disorders, and conditions associated with CHL2 polypeptides.

## Background of the Invention

Technical advances in the identification, cloning, expression, and manipulation of nucleic acid molecules and the deciphering of the human genome have greatly accelerated the discovery of novel therapeutics. Rapid nucleic acid sequencing techniques can now generate sequence information at unprecedented rates and, coupled with computational analyses, allow the assembly of overlapping sequences into partial and entire genomes and the identification of polypeptide-encoding regions. A comparison of a predicted amino acid sequence against a database compilation of known amino acid sequences allows one to determine the extent of homology to previously identified sequences and/or structural landmarks. The cloning and expression of a polypeptide-encoding region of a nucleic acid molecule provides a polypeptide product for structural and functional analyses. The manipulation of nucleic acid molecules and encoded polypeptides may confer advantageous properties on a product for use as a therapeutic.

In spite of the significant technical advances in genome research over the past decade, the potential for the development of novel therapeutics based on the human genome is still largely unrealized. Many genes encoding potentially beneficial polypeptide therapeutics or those encoding polypeptides, which may act as "targets" for therapeutic molecules, have still not been identified.

Accordingly, it is an object of the invention to identify novel polypeptides,

and nucleic acid molecules encoding the same, which have diagnostic or therapeutic benefit.

CHL2 is structurally related to the bone morphogenetic protein (BMP) inhibitor known as chordin (CHD), (Sasai et al., 1994, *Cell* 79:779-90), or short gastrulation (SOG; Francois, et al., 1994, *Genes Dev.* 8:2602-16). The CHL2 gene is believed to be a member of CHD/SOG family.

Bone morphogenetic protein (BMP) is a member of the transforming growth factor–beta family, which was originally identified as a factor promoting bone formation from a cartridge implant (Wozney *et al.*, 1988, *Science* 242:1528-34; Celeste *et al.*, 1990, *Proc. Nat. Acad. Sci. USA* 87:9843-47). BMP is also known to play an essential role during the early embryogenesis of the frog, the fly, and in mammals. The precise concentration of active BMP seems to be important for the specification of particular cell types (Dale *et al.*, 1992, *Development* 115:573-85; Dosch *et al.*, 1997, *Development* 124:2325-34). An activity gradient of BMP2/4 is observed in, for example, *Xenopus* embryos in which the lowest expression is detected at the dorsal tip and the highest expression at the ventral tip – establishing the dorsoventral axis determination in the embryo. In another example, the control of BMP concentration at specific sites of tissue development suggests a role for BMP in organogenesis. Control of BMP expression is achieved by either localized expression of the BMP gene products or through the influence of the BMP inhibitor chordin (CHD) (Sasai *et al.*, 1994, *Cell* 79:779-90) – or short gastrulation (SOG) (Francois *et al.*, 1994, *Genes Dev.* 8:2602-16).

CHD/SOG is a large secreted protein produced from the Spemann's organizer, the master-controlling region for the dorsoventral axis specification at the gastrulation stage of *Xenopus* embryogenesis. CHD/SOG functions as a dorsalization factor, as does Noggin (Smith and Harland, 1992, *Cell* 70:829-40), which is also secreted from the organizer. The *Drosophila* SOG has a transmembrane domain at its amino-terminus, suggesting that it may be a type II transmembrane protein (Francois *et al.*, 1994, *Genes Dev.* 8:2602-16). It has been proposed that the carboxyl-terminal side (extracellular domain) of the *Drosophila*

SOG is cleaved off. However, *Xenopus* CHD (Sasai *et al.*, 1994, *Cell* 79:779-90), Zebrafish CHD (Schulte-Merker *et al.*, 1997, *Nature* 387:862-63), and murine CHD (Pappano *et al.*, 1998, *Genomics* 52:236-39) do not contain the transmembrane domain. Instead, these proteins have a signal peptide, and are secreted directly. The CHD/SOG polypeptide contains four repeats of the cysteine-rich domain (CR1-4) that is also found in a variety of extracellular matrix proteins such as collagen and thrombospondin.

CHD/SOG is known to bind to one of the ventralizing factors, BMP4 (Piccolo *et al.*, 1996, *Cell* 86:589-98). BMP4 has been shown to be essential for embryonic development of posterior-ventral mesoderm in mice (Winnier *et al.*, 1995, *Genes Dev.* 9:2105-16). The binding of CHD/SOG to BMP4 inhibits BMP4 activity by preventing BMP4 from binding to its receptor (Piccolo *et al.*, 1996, *Cell* 86:589-98). In this respect, the functional relationship between CHD/SOG and BMP4 resembles that between OPG and OPGL, although CHD/SOG is not structurally related to the BMP receptors. The binding affinity of CHD/SOG to BMP4 is specific and tight (Kd = 3 x $10^{-10}$ M (Piccolo *et al.*, 1996, *Cell* 86:589-98), and seems to require proteolysis in order to effectuate the release of bound BMP4. This proteolysis is achieved by a specific metalloprotease – Tolloid (TLD) or BMP1 – that cleaves CHD/SOG to liberate either, or both, the first (CR1) and last (CR4) CR motifs (Piccolo *et al.*, 1997, *Cell* 91: 407-16). Whether or not CHD/SOG has other functions or an independent function through its own receptor remains to be determined.

One of the most important roles of CHD/SOG is to establish a BMP4 morphogen gradient (Jones and Smith, 1998, *Dev. Biol.* 194:12-17). BMP4 itself only migrates a short distance and seems to act essentially on the cell autonomously (Jones *et al.*, 1996, *Curr. Biol.* 6:1468-75). In contrast, the BMP4 inhibitors Noggin and CHD/SOG appear to exert a long-range effect, thereby forming an activity gradient of BMP4.

BMPs also play important roles outside of early embryogenesis, for example in the organogenesis of lung, gut, kidney, skin, heart and teeth, as well as in the later stages of embryogenesis (Hogan, 1996, *Genes Dev.* 10:1580-94).

Some BMPs are expressed in a very localized fashion while others are expressed widely in a tissue. The importance of the localized action of BMP for organogenesis has been supported by transgenic mouse experiments using constructs by which BMP concentration is artificially elevated throughout the target tissue. In the case of lung, BMP4 is expressed in the distal tips of epithelium in the developing lung, and when overexpressed with the surfactant protein C promoter, the development of a small lung in which the structural organization (*i.e.*, branching) has been severely disrupted is observed (Bellusci *et al.*, 1996, *Development* 122:1693-702). Since the putative BMP-activity gradient could also be disrupted by the transgene expression, BMPs expressed widely in the tissue could also play a role in the determination of the structural organization of a tissue.

Noggin is another BMP2/4 inhibitor secreted from Spemann's organizer (Zimmerman *et al.*, 1996, *Cell* 86:599-606). The biological role of Noggin and its mode of action are similar to CHD/SOG in *Xenopus*. Although the most notable function of Noggin is, like CHD/SOG, dorsalization, Noggin null-mutant mice have shown a bone phenotype (hyperplasia of chondrocytes) instead of an early embryonic phenotype (McMahon *et al.*, 1998, *Genes Dev.* 12:1438-52; Brunet *et al.*, 1998, *Science* 280:1455-57). This suggests that CHL2 or even CHD might have a non-dispensable function in the later stage of embryogenesis.

Summary of the Invention

The present invention relates to novel CHL2 nucleic acid molecules and encoded polypeptides.

The invention provides for an isolated nucleic acid molecule comprising a nucleotide sequence selected from the group consisting of:

(a) the nucleotide sequence as set forth in either SEQ ID NO: 1 or SEQ ID NO: 4;

(b) the nucleotide sequence of the DNA insert either ATCC Deposit Nos. PTA-1479 or PTA-1480;

(c) a nucleotide sequence encoding the polypeptide as set forth in either SEQ ID NO: 2 or SEQ ID NO: 5;

(d) a nucleotide sequence which hybridizes under moderately or highly stringent conditions to the complement of any of (a) - (c); and

(e) a nucleotide sequence complementary to any of (a) - (c).

The invention also provides for an isolated nucleic acid molecule comprising a nucleotide sequence selected from the group consisting of:

(a) a nucleotide sequence encoding a polypeptide which is at least about 70 percent identical to the polypeptide as set forth in either SEQ ID NO: 2 or SEQ ID NO: 5, wherein the encoded polypeptide has an activity of the polypeptide set forth in either SEQ ID NO: 2 or SEQ ID NO: 5;

(b) a nucleotide sequence encoding an allelic variant or splice variant of the nucleotide sequence as set forth in either SEQ ID NO: 1 or SEQ ID NO: 4, the nucleotide sequence of the DNA insert in either ATCC Deposit Nos. PTA-1479 or PTA-1480, or (a);

(c) a region of the nucleotide sequence of any of SEQ ID NO: 1 or SEQ ID NO: 4, the DNA insert in either ATCC Deposit Nos. PTA-1479 or PTA-1480, (a), or (b) encoding a polypeptide fragment of at least about 25 amino acid residues, wherein the polypeptide fragment has an activity of the polypeptide set forth in either SEQ ID NO: 2 or SEQ ID NO: 5, or is antigenic;

(d) a region of the nucleotide sequence of any of SEQ ID NO: 1 or SEQ ID NO: 4, the DNA insert in either ATCC Deposit Nos. PTA-1479 or PTA-1480, or any of (a) - (c) comprising a fragment of at least about 16 nucleotides;

(e) a nucleotide sequence which hybridizes under moderately or highly stringent conditions to the complement of any of (a) - (d); and

(f) a nucleotide sequence complementary to any of (a) - (d).

The invention further provides for an isolated nucleic acid molecule comprising a nucleotide sequence selected from the group consisting of:

(a) a nucleotide sequence encoding a polypeptide as set forth in either SEQ ID NO: 2 or SEQ ID NO: 5 with at least one conservative amino acid substitution, wherein the encoded polypeptide has an activity of the polypeptide set forth in either SEQ ID NO: 2 or SEQ ID NO: 5;

(b) a nucleotide sequence encoding a polypeptide as set forth in either SEQ ID NO: 2 or SEQ ID NO: 5 with at least one amino acid insertion, wherein the encoded polypeptide has an activity of the polypeptide set forth in either SEQ ID NO: 2 or SEQ ID NO: 5;

(c) a nucleotide sequence encoding a polypeptide as set forth in either SEQ ID NO: 2 or SEQ ID NO: 5 with at least one amino acid deletion, wherein the encoded polypeptide has an activity of the polypeptide set forth in either SEQ ID NO: 2 or SEQ ID NO: 5;

(d) a nucleotide sequence encoding a polypeptide as set forth in either SEQ ID NO: 2 or SEQ ID NO: 5 which has a C- and/or N- terminal truncation, wherein the encoded polypeptide has an activity of the polypeptide set forth in either SEQ ID NO: 2 or SEQ ID NO: 5;

(e) a nucleotide sequence encoding a polypeptide as set forth in either SEQ ID NO: 2 or SEQ ID NO: 5 with at least one modification selected from the group consisting of amino acid substitutions, amino acid insertions, amino acid deletions, C-terminal truncation, and N-terminal truncation, wherein the encoded polypeptide has an activity of the polypeptide set forth in either SEQ ID NO: 2 or SEQ ID NO: 5;

(f) a nucleotide sequence of any of (a) - (e) comprising a fragment of at least about 16 nucleotides;

(g) a nucleotide sequence which hybridizes under moderately or highly stringent conditions to the complement of any of (a) - (f); and

(h) a nucleotide sequence complementary to any of (a) - (e).

The present invention provides for an isolated polypeptide comprising an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence as set forth in either SEQ ID NO: 2 or SEQ ID NO: 5; and

(b) the amino acid sequence encoded by the DNA insert in either ATCC Deposit Nos. PTA-1479 or PTA-1480.

The invention also provides for an isolated polypeptide comprising the amino acid sequence selected from the group consisting of:

(a) the amino acid sequence as set forth in either SEQ ID NO: 3 or SEQ ID NO: 6, optionally further comprising an amino-terminal methionine;

(b) an amino acid sequence for an ortholog of any of SEQ ID NO: 2 or SEQ ID NO: 5;

(c) an amino acid sequence which is at least about 70 percent identical to the amino acid sequence of any of SEQ ID NO: 2 or SEQ ID NO: 5, wherein the polypeptide has an activity of the polypeptide set forth in either SEQ ID NO: 2 or SEQ ID NO: 5;

(d) a fragment of the amino acid sequence set forth in either SEQ ID NO: 2 or SEQ ID NO: 5 comprising at least about 25 amino acid residues, wherein the fragment has an activity of the polypeptide set forth in either SEQ ID NO: 2 or SEQ ID NO: 5, or is antigenic; and

(e) an amino acid sequence for an allelic variant or splice variant of the amino acid sequence as set forth in either SEQ ID NO: 2 or SEQ ID NO: 5, the amino acid sequence encoded by the DNA insert in either ATCC Deposit Nos. PTA-1479 or PTA-1480, or any of (a) - (c).

The invention further provides for an isolated polypeptide comprising the amino acid sequence selected from the group consisting of:

(a) the amino acid sequence as set forth in either SEQ ID NO: 2 or SEQ ID NO: 5 with at least one conservative amino acid substitution, wherein the polypeptide has an activity of the polypeptide set forth in either SEQ ID NO: 2 or SEQ ID NO: 5;

(b) the amino acid sequence as set forth in either SEQ ID NO: 2 or SEQ ID NO: 5 with at least one amino acid insertion, wherein the polypeptide has an activity of the polypeptide set forth in either SEQ ID NO: 2 or SEQ ID NO: 5;

(c) the amino acid sequence as set forth in either SEQ ID NO: 2 or SEQ ID NO: 5 with at least one amino acid deletion, wherein the polypeptide has an activity of the polypeptide set forth in either SEQ ID NO: 2 or SEQ ID NO: 5;

(d) the amino acid sequence as set forth in either SEQ ID NO: 2 or SEQ ID NO: 5 which has a C- and/or N-terminal truncation, wherein the polypeptide has an activity of the polypeptide set forth in either SEQ ID NO: 2 or SEQ ID NO: 5; and

(e) the amino acid sequence as set forth in either SEQ ID NO: 2 or SEQ ID NO: 5 with at least one modification selected from the group consisting of amino acid substitutions, amino acid insertions, amino acid deletions, C-terminal truncation, and N-terminal truncation, wherein the polypeptide has an activity of the polypeptide set forth in either SEQ ID NO: 2 or SEQ ID NO: 5.

The invention still further provides for an isolated polypeptide comprising the amino acid sequence as set forth in SEQ ID NO: 5 with at least one conservative amino acid substitution selected from the group consisting of: leucine or methionine at position 2; methionine at position 5; lysine at position 6; alanine at position 7; isoleucine at position 8; phenylalanine at position 14; leucine at position 15; threonine at position 23; leucine at position 25; valine at position 27; glutamic acid at position 30; tyrosine at position 32; methionine at position 34; glutamine at position 36; lysine at position 39; alanine at position 41; threonine at position 45; valine at position 55; valine at position 59; asparagine at position 60; proline at position 66; asparagine at position 68; serine or threonine at position 72; valine at position 74; arginine at position 75; arginine at position 94; asparagine at position 99; serine at position 100; lysine at position 105; valine at position 106; tyrosine at position 113; serine at position 116; serine at position 118; arginine at position 120; leucine at position 123; alanine at position 125; alanine at position 129; alanine at position 139; threonine at position 142; serine

at position 144; asparagine at position 147; valine at position 148; serine at position 149; alanine at position 159; alanine at position 160; alanine at position 161; valine at position 164; valine at position 166; valine at position 173; arginine at position 175; aspartic acid at position 177; alanine at position 190; phenylalanine at position 191; arginine at position 192; leucine at position 194; asparagine at position 196; leucine at position 205; alanine at position 210; alanine at position 212; serine at position 213; alanine at position 216; serine at position 217; alanine at position 218; isoleucine at position 219; alanine at position 222; leucine at position 225; phenylalanine at position 226; leucine at position 230; glutamine or arginine at position 233; glutamine at position 241; leucine at position 242; isoleucine at position 244; glutamine or asparagine at position 245; glutamine at position 249; leucine or valine at position 251; alanine at position 256; asparagine at position 257; serine at position 259; alanine at position 260; glutamine at position 261; phenylalanine at position 265; valine at position 268; leucine at position 269; leucine at position 272; valine at position 275; valine at position 278; glutamic acid at position 284; glutamic acid at position 288; alanine or isoleucine at position 292; serine at position 300; isoleucine at position 306; valine at position 313; serine at position 314; leucine at position 319; glutamine at position 323; threonine at position 324; alanine at position 326; alanine at position 327; serine at position 329; serine at position 331; leucine at position 334; asparagine at position 337; valine at position 339; leucine at position 340; serine at position 342; phenylalanine at position 344; glutamic acid at position 349; isoleucine or valine at position 354; methionine at position 356; valine at position 366; methionine or valine at position 368; isoleucine at position 371; leucine at position 375; leucine at position 376; glutamine at position 377; phenylalanine at position 381; asparagine at position 383; isoleucine at position 384; leucine at position 393; arginine at position 395; valine at position 398; alanine or valine at position 399; tyrosine at position 403; asparagine at position 406; isoleucine at position 409; alanine or valine at position 415; isoleucine at position 417; and leucine at position 421; wherein the polypeptide has an activity of the polypeptide as set forth in SEQ ID NO: 5.

Also provided are fusion polypeptides comprising CHL2 amino acid sequences.

The present invention also provides for an expression vector comprising the isolated nucleic acid molecules as set forth herein, recombinant host cells comprising the recombinant nucleic acid molecules as set forth herein, and a method of producing a CHL2 polypeptide comprising culturing the host cells and optionally isolating the polypeptide so produced.

A transgenic non-human animal comprising a nucleic acid molecule encoding a CHL2 polypeptide is also encompassed by the invention. The CHL2 nucleic acid molecules are introduced into the animal in a manner that allows expression and increased levels of a CHL2 polypeptide, which may include increased circulating levels. Alternatively, the CHL2 nucleic acid molecules are introduced into the animal in a manner that prevents expression of endogenous CHL2 polypeptide (*i.e.*, generates a transgenic animal possessing a CHL2 polypeptide gene knockout). The transgenic non-human animal is preferably a mammal, and more preferably a rodent, such as a rat or a mouse.

Also provided are derivatives of the CHL2 polypeptides of the present invention.

Additionally provided are selective binding agents such as antibodies and peptides capable of specifically binding the CHL2 polypeptides of the invention. Such antibodies and peptides may be agonistic or antagonistic.

Pharmaceutical compositions comprising the nucleotides, polypeptides, or selective binding agents of the invention and one or more pharmaceutically acceptable formulation agents are also encompassed by the invention. The pharmaceutical compositions are used to provide therapeutically effective amounts of the nucleotides or polypeptides of the present invention. The invention is also directed to methods of using the polypeptides, nucleic acid molecules, and selective binding agents.

The CHL2 polypeptides and nucleic acid molecules of the present invention may be used to treat, prevent, ameliorate, and/or detect diseases and disorders, including those recited herein.

The present invention also provides a method of assaying test molecules to identify a test molecule that binds to a CHL2 polypeptide. The method comprises contacting a CHL2 polypeptide with a test molecule to determine the extent of binding of the test molecule to the polypeptide. The method further comprises determining whether such test molecules are agonists or antagonists of a CHL2 polypeptide. The present invention further provides a method of testing the impact of molecules on the expression of CHL2 polypeptide or on the activity of CHL2 polypeptide.

Methods of regulating expression and modulating (*i.e.*, increasing or decreasing) levels of a CHL2 polypeptide are also encompassed by the invention. One method comprises administering to an animal a nucleic acid molecule encoding a CHL2 polypeptide. In another method, a nucleic acid molecule comprising elements that regulate or modulate the expression of a CHL2 polypeptide may be administered. Examples of these methods include gene therapy, cell therapy, and anti-sense therapy as further described herein.

In another aspect of the present invention, the CHL2 polypeptides may be used for identifying receptors thereof ("CHL2 polypeptide receptors"). Various forms of "expression cloning" have been extensively used to clone receptors for protein ligands. *See, e.g.*, Simonsen and Lodish, 1994, *Trends Pharmacol. Sci.* 15:437-41 and Tartaglia *et al.*, 1995, *Cell* 83:1263-71. The isolation of a CHL2 polypeptide receptor is useful for identifying or developing novel agonists and antagonists of the CHL2 polypeptide signaling pathway. Such agonists and antagonists include soluble CHL2 polypeptide receptors, anti-CHL2 polypeptide receptor-selective binding agents (such as antibodies and derivatives thereof), small molecules, and antisense oligonucleotides, any of which can be used for treating one or more disease or disorder, including those disclosed herein.

Brief Description of the Figures

Figures 1A-1C illustrate the nucleotide sequence of the murine CHL2 gene (SEQ ID NO: 1) and the deduced amino acid sequence of murine CHL2 polypeptide (SEQ ID NO: 2). The predicted signal sequence is indicated (underline);

Figure 2 illustrates an amino acid sequence alignment of murine CHL2 polypeptide (mouse CHL2; SEQ ID NO: 2) and murine chordin (Af069501; SEQ ID NO: 7);

Figures 3A-3C illustrate the nucleotide sequence of the human CHL2 gene (SEQ ID NO: 4) and the deduced amino acid sequence of human CHL2 polypeptide (SEQ ID NO: 5). The predicted signal sequence is indicated (underline);

Figure 4 illustrates an amino acid sequence alignment of human chordin (huCHD; SEQ ID NO: 8), human CHL1 polypeptide (huCHL; SEQ ID NO: 9), and human CHL2 polypeptide (huCHL2; SEQ ID NO: 5);

Figure 5 illustrates a schematic representation of murine chordin (Chordin), CHL1 polypeptide, and CHL2 polypeptide. Pro-collagen repeats (CR; homologous CR domains are indicated by gray boxes), signal peptides (SP), putative BMP1/Tolloid cleavage sites (*), and sites of amino acid sequence variation in CHL1 (dE and d5) are indicated;

Figure 6 illustrates the expression of murine CHL2 mRNA as detected by *in situ* hybridization in E17.5 mouse hip joint (panels A and B) and costal-chondral articulation (panel C; showing signal in articular chondrocytes on both sides of the articulation);

Figure 7 illustrates the expression of murine CHL2 mRNA as detected by *in situ* hybridization in normal adult vertebral articulation (panels A, B, and C; showing signal in articular chondrocytes at the surface of the articular cartilage on both sides of the zygapophyseal or facet joint between the processes of adjacent

vertebrae) and the fibrocartilage of the anulus fibrosus of a vertebral disc (panel D);

Figure 8 illustrates the expression of murine CHL2 mRNA as detected by *in situ* hybridization in E18.5 mouse sternum and placenta and normal adult mouse uterus, colon, and small intestine;

Figure 9 illustrates the secondary axis-forming activity of murine CHL2 polypeptide;

Figure 10 illustrates the effect of CHL2 polypeptide on the BMP4-dependent generation of $CD34^+/CD31^+$ erythro-myeloid progenitor cells (R7) and CD34-/CD31+ cells (R3);

Figure 11 illustrates the effect of CHL2 polypeptide on the BMP2-dependent induction of alkaline phosphatase in C2C12 myoblastic cells;

Figure 12 illustrates the results of Western blot analysis using mCHL2-FLAG proteins. In panel A, mCHL2-FLAG was mixed with BMP5 or activin A, treated with an anti-mCHL2 antiserum, and precipitated with protein A agarose beads. In panel B, mCHL2-FLAG was mixed with BMP2, BMP4, BMP5, BMP6, GDF5 (BMP14), or activin A, treated with an anti-mCHL2 antiserum, and precipitated with protein A agarose beads.

Detailed Description of the Invention

The section headings used herein are for organizational purposes only and are not to be construed as limiting the subject matter described. All references cited in this application are expressly incorporated by reference herein.

Definitions

The terms "CHL2 gene" or "CHL2 nucleic acid molecule" or "CHL2 polynucleotide" refer to a nucleic acid molecule comprising or consisting of a nucleotide sequence as set forth in either SEQ ID NO: 1 or SEQ ID NO: 4, a nucleotide sequence encoding the polypeptide as set forth in either SEQ ID NO: 2 or SEQ ID NO: 5, a nucleotide sequence of the DNA insert in either ATCC Deposit Nos. PTA-1479 or PTA-1480, and nucleic acid molecules as defined herein.

The term "CHL2 polypeptide allelic variant" refers to one of several possible naturally occurring alternate forms of a gene occupying a given locus on a chromosome of an organism or a population of organisms.

The term "CHL2 polypeptide splice variant" refers to a nucleic acid molecule, usually RNA, which is generated by alternative processing of intron sequences in an RNA transcript of CHL2 polypeptide amino acid sequence as set forth in either SEQ ID NO: 2 or SEQ ID NO: 5.

The term "isolated nucleic acid molecule" refers to a nucleic acid molecule of the invention that (1) has been separated from at least about 50 percent of proteins, lipids, carbohydrates, or other materials with which it is naturally found when total nucleic acid is isolated from the source cells, (2) is not linked to all or a portion of a polynucleotide to which the "isolated nucleic acid molecule" is linked in nature, (3) is operably linked to a polynucleotide which it is not linked to in nature, or (4) does not occur in nature as part of a larger polynucleotide sequence. Preferably, the isolated nucleic acid molecule of the present invention is substantially free from any other contaminating nucleic acid molecule(s) or other contaminants that are found in its natural environment that would interfere with its use in polypeptide production or its therapeutic, diagnostic, prophylactic or research use.

The term "nucleic acid sequence" or "nucleic acid molecule" refers to a DNA or RNA sequence. The term encompasses molecules formed from any of the known base analogs of DNA and RNA such as, but not limited to 4-acetylcytosine, 8-hydroxy-N6-methyladenosine, aziridinyl-cytosine, pseudoisocytosine, 5-(carboxyhydroxylmethyl) uracil, 5-fluorouracil, 5-

bromouracil, 5-carboxymethylaminomethyl-2-thiouracil, 5-carboxymethylaminomethyluracil, dihydrouracil, inosine, N6-iso-pentenyladenine, 1-methyladenine, 1-methylpseudouracil, 1-methylguanine, 1-methylinosine, 2,2-dimethyl-guanine, 2-methyladenine, 2-methylguanine, 3-methylcytosine, 5-methylcytosine, N6-methyladenine, 7-methylguanine, 5-methylaminomethyluracil, 5-methoxyamino-methyl-2-thiouracil, beta-D-mannosylqueosine, 5' -methoxycarbonyl-methyluracil, 5-methoxyuracil, 2-methylthio-N6-isopentenyladenine, uracil-5-oxyacetic acid methylester, uracil-5-oxyacetic acid, oxybutoxosine, pseudouracil, queosine, 2-thiocytosine, 5-methyl-2-thiouracil, 2-thiouracil, 4-thiouracil, 5-methyluracil, N-uracil-5-oxyacetic acid methylester, uracil-5-oxyacetic acid, pseudouracil, queosine, 2-thiocytosine, and 2,6-diaminopurine.

The term "vector" is used to refer to any molecule (*e.g.*, nucleic acid, plasmid, or virus) used to transfer coding information to a host cell.

The term "expression vector" refers to a vector that is suitable for transformation of a host cell and contains nucleic acid sequences that direct and/or control the expression of inserted heterologous nucleic acid sequences. Expression includes, but is not limited to, processes such as transcription, translation, and RNA splicing, if introns are present.

The term "operably linked" is used herein to refer to an arrangement of flanking sequences wherein the flanking sequences so described are configured or assembled so as to perform their usual function. Thus, a flanking sequence operably linked to a coding sequence may be capable of effecting the replication, transcription and/or translation of the coding sequence. For example, a coding sequence is operably linked to a promoter when the promoter is capable of directing transcription of that coding sequence. A flanking sequence need not be contiguous with the coding sequence, so long as it functions correctly. Thus, for example, intervening untranslated yet transcribed sequences can be present between a promoter sequence and the coding sequence and the promoter sequence can still be considered "operably linked" to the coding sequence.

The term "host cell" is used to refer to a cell which has been transformed, or is capable of being transformed with a nucleic acid sequence and then of expressing a selected gene of interest. The term includes the progeny of the parent cell, whether or not the progeny is identical in morphology or in genetic make-up to the original parent, so long as the selected gene is present.

The term "CHL2 polypeptide" refers to a polypeptide comprising the amino acid sequence of any of SEQ ID NO: 2 or SEQ ID NO: 5 and related polypeptides. Related polypeptides include CHL2 polypeptide fragments, CHL2 polypeptide orthologs, CHL2 polypeptide variants, and CHL2 polypeptide derivatives, which possess at least one activity of the polypeptide as set forth in either SEQ ID NO: 2 or SEQ ID NO: 5. CHL2 polypeptides may be mature polypeptides, as defined herein, and may or may not have an amino-terminal methionine residue, depending on the method by which they are prepared.

The term "CHL2 polypeptide fragment" refers to a polypeptide that comprises a truncation at the amino-terminus (with or without a leader sequence) and/or a truncation at the carboxyl-terminus of the polypeptide as set forth in either SEQ ID NO: 2 or SEQ ID NO: 5. The term "CHL2 polypeptide fragment" also refers to amino-terminal and/or carboxyl-terminal truncations of CHL2 polypeptide orthologs, CHL2 polypeptide derivatives, or CHL2 polypeptide variants, or to amino-terminal and/or carboxyl-terminal truncations of the polypeptides encoded by CHL2 polypeptide allelic variants or CHL2 polypeptide splice variants. CHL2 polypeptide fragments may result from alternative RNA splicing or from *in vivo* protease activity. Membrane-bound forms of a CHL2 polypeptide are also contemplated by the present invention. In preferred embodiments, truncations and/or deletions comprise about 10 amino acids, or about 20 amino acids, or about 50 amino acids, or about 75 amino acids, or about 100 amino acids, or more than about 100 amino acids. The polypeptide fragments so produced will comprise about 25 contiguous amino acids, or about 50 amino acids, or about 75 amino acids, or about 100 amino acids, or about 150 amino acids, or about 200 amino acids, or more than about 200 amino acids. Such CHL2 polypeptide fragments may optionally comprise an amino-terminal methionine

residue. It will be appreciated that such fragments can be used, for example, to generate antibodies to CHL2 polypeptides.

The term "CHL2 polypeptide ortholog" refers to a polypeptide from another species that corresponds to CHL2 polypeptide amino acid sequence as set forth in either SEQ ID NO: 2 or SEQ ID NO: 5. For example, mouse and human CHL2 polypeptides are considered orthologs of each other.

The term "CHL2 polypeptide variants" refers to CHL2 polypeptides comprising amino acid sequences having one or more amino acid sequence substitutions, deletions (such as internal deletions and/or CHL2 polypeptide fragments), and/or additions (such as internal additions and/or CHL2 fusion polypeptides) as compared to the CHL2 polypeptide amino acid sequence set forth in either SEQ ID NO: 2 or SEQ ID NO: 5 (with or without a leader sequence). Variants may be naturally occurring (*e.g.*, CHL2 polypeptide allelic variants, CHL2 polypeptide orthologs, and CHL2 polypeptide splice variants) or artificially constructed. Such CHL2 polypeptide variants may be prepared from the corresponding nucleic acid molecules having a DNA sequence that varies accordingly from the DNA sequence as set forth in either SEQ ID NO: 1 or SEQ ID NO: 4. In preferred embodiments, the variants have from 1 to 3, or from 1 to 5, or from 1 to 10, or from 1 to 15, or from 1 to 20, or from 1 to 25, or from 1 to 50, or from 1 to 75, or from 1 to 100, or more than 100 amino acid substitutions, insertions, additions and/or deletions, wherein the substitutions may be conservative, or non-conservative, or any combination thereof.

The term "CHL2 polypeptide derivatives" refers to the polypeptide as set forth in either SEQ ID NO: 2 or SEQ ID NO: 5, CHL2 polypeptide fragments, CHL2 polypeptide orthologs, or CHL2 polypeptide variants, as defined herein, that have been chemically modified. The term "CHL2 polypeptide derivatives" also refers to the polypeptides encoded by CHL2 polypeptide allelic variants or CHL2 polypeptide splice variants, as defined herein, that have been chemically modified.

The term "mature CHL2 polypeptide" refers to a CHL2 polypeptide lacking a leader sequence. A mature CHL2 polypeptide may also include other

modifications such as proteolytic processing of the amino-terminus (with or without a leader sequence) and/or the carboxyl-terminus, cleavage of a smaller polypeptide from a larger precursor, N-linked and/or O-linked glycosylation, and the like. Exemplary mature CHL2 polypeptides are depicted by the amino acid sequences of SEQ ID NO: 3 and SEQ ID NO: 6.

The term "CHL2 fusion polypeptide" refers to a fusion of one or more amino acids (such as a heterologous protein or peptide) at the amino- or carboxyl-terminus of the polypeptide as set forth in either SEQ ID NO: 2 or SEQ ID NO: 5, CHL2 polypeptide fragments, CHL2 polypeptide orthologs, CHL2 polypeptide variants, or CHL2 derivatives, as defined herein. The term "CHL2 fusion polypeptide" also refers to a fusion of one or more amino acids at the amino- or carboxyl-terminus of the polypeptide encoded by CHL2 polypeptide allelic variants or CHL2 polypeptide splice variants, as defined herein.

The term "biologically active CHL2 polypeptides" refers to CHL2 polypeptides having at least one activity characteristic of the polypeptide comprising the amino acid sequence of any of SEQ ID NO: 2 or SEQ ID NO: 5. In addition, a CHL2 polypeptide may be active as an immunogen; that is, the CHL2 polypeptide contains at least one epitope to which antibodies may be raised.

The term "isolated polypeptide" refers to a polypeptide of the present invention that (1) has been separated from at least about 50 percent of polynucleotides, lipids, carbohydrates, or other materials with which it is naturally found when isolated from the source cell, (2) is not linked (by covalent or noncovalent interaction) to all or a portion of a polypeptide to which the "isolated polypeptide" is linked in nature, (3) is operably linked (by covalent or noncovalent interaction) to a polypeptide with which it is not linked in nature, or (4) does not occur in nature. Preferably, the isolated polypeptide is substantially free from any other contaminating polypeptides or other contaminants that are found in its natural environment that would interfere with its therapeutic, diagnostic, prophylactic or research use.

The term "identity," as known in the art, refers to a relationship between the sequences of two or more polypeptide molecules or two or more nucleic acid molecules, as determined by comparing the sequences. In the art, "identity" also means the degree of sequence relatedness between nucleic acid molecules or polypeptides, as the case may be, as determined by the match between strings of two or more nucleotide or two or more amino acid sequences. "Identity" measures the percent of identical matches between the smaller of two or more sequences with gap alignments (if any) addressed by a particular mathematical model or computer program (*i.e.*, "algorithms").

The term "similarity" is a related concept, but in contrast to "identity," "similarity" refers to a measure of relatedness which includes both identical matches and conservative substitution matches. If two polypeptide sequences have, for example, 10/20 identical amino acids, and the remainder are all non-conservative substitutions, then the percent identity and similarity would both be 50%. If in the same example, there are five more positions where there are conservative substitutions, then the percent identity remains 50%, but the percent similarity would be 75% (15/20). Therefore, in cases where there are conservative substitutions, the percent similarity between two polypeptides will be higher than the percent identity between those two polypeptides.

The term "naturally occurring" or "native" when used in connection with biological materials such as nucleic acid molecules, polypeptides, host cells, and the like, refers to materials which are found in nature and are not manipulated by man. Similarly, "non-naturally occurring" or "non-native" as used herein refers to a material that is not found in nature or that has been structurally modified or synthesized by man.

The terms "effective amount" and "therapeutically effective amount" each refer to the amount of a CHL2 polypeptide or CHL2 nucleic acid molecule used to support an observable level of one or more biological activities of the CHL2 polypeptides as set forth herein.

The term "pharmaceutically acceptable carrier" or "physiologically acceptable carrier" as used herein refers to one or more formulation materials

suitable for accomplishing or enhancing the delivery of the CHL2 polypeptide, CHL2 nucleic acid molecule, or CHL2 selective binding agent as a pharmaceutical composition.

The term "antigen" refers to a molecule or a portion of a molecule capable of being bound by a selective binding agent, such as an antibody, and additionally capable of being used in an animal to produce antibodies capable of binding to an epitope of that antigen. An antigen may have one or more epitopes.

The term "selective binding agent" refers to a molecule or molecules having specificity for a CHL2 polypeptide. As used herein, the terms, "specific" and "specificity" refer to the ability of the selective binding agents to bind to human CHL2 polypeptides and not to bind to human non-CHL2 polypeptides. It will be appreciated, however, that the selective binding agents may also bind orthologs of the polypeptide as set forth in either SEQ ID NO: 2 or SEQ ID NO: 5, that is, interspecies versions thereof, such as mouse and rat CHL2 polypeptides.

The term "transduction" is used to refer to the transfer of genes from one bacterium to another, usually by a phage. "Transduction" also refers to the acquisition and transfer of eukaryotic cellular sequences by retroviruses.

The term "transfection" is used to refer to the uptake of foreign or exogenous DNA by a cell, and a cell has been "transfected" when the exogenous DNA has been introduced inside the cell membrane. A number of transfection techniques are well known in the art and are disclosed herein. *See, e.g.*, Graham *et al.*, 1973, *Virology* 52:456; Sambrook *et al.*, *Molecular Cloning, A Laboratory Manual* (Cold Spring Harbor Laboratories, 1989); Davis *et al.*, *Basic Methods in Molecular Biology* (Elsevier, 1986); and Chu *et al.*, 1981, *Gene* 13:197. Such techniques can be used to introduce one or more exogenous DNA moieties into suitable host cells.

The term "transformation" as used herein refers to a change in a cell's genetic characteristics, and a cell has been transformed when it has been modified to contain a new DNA. For example, a cell is transformed where it is genetically modified from its native state. Following transfection or transduction, the transforming DNA may recombine with that of the cell by physically integrating

into a chromosome of the cell, may be maintained transiently as an episomal element without being replicated, or may replicate independently as a plasmid. A cell is considered to have been stably transformed when the DNA is replicated with the division of the cell.

Relatedness of Nucleic Acid Molecules and/or Polypeptides

It is understood that related nucleic acid molecules include allelic or splice variants of the nucleic acid molecule of any of SEQ ID NO: 1 or SEQ ID NO: 4, and include sequences which are complementary to any of the above nucleotide sequences. Related nucleic acid molecules also include a nucleotide sequence encoding a polypeptide comprising or consisting essentially of a substitution, modification, addition and/or deletion of one or more amino acid residues compared to the polypeptide in either SEQ ID NO: 2 or SEQ ID NO: 5. Such related CHL2 polypeptides may comprise, for example, an addition and/or a deletion of one or more N-linked or O-linked glycosylation sites or an addition and/or a deletion of one or more cysteine residues.

Related nucleic acid molecules also include fragments of CHL2 nucleic acid molecules which encode a polypeptide of at least about 25 contiguous amino acids, or about 50 amino acids, or about 75 amino acids, or about 100 amino acids, or about 150 amino acids, or about 200 amino acids, or more than 200 amino acid residues of the CHL2 polypeptide of any of SEQ ID NO: 2 or SEQ ID NO: 5.

In addition, related CHL2 nucleic acid molecules also include those molecules which comprise nucleotide sequences which hybridize under moderately or highly stringent conditions as defined herein with the fully complementary sequence of the CHL2 nucleic acid molecule of any of SEQ ID NO: 1 or SEQ ID NO: 4, or of a molecule encoding a polypeptide, which polypeptide comprises the amino acid sequence as shown in either SEQ ID NO: 2 or SEQ ID NO: 5, or of a nucleic acid fragment as defined herein, or of a nucleic acid fragment encoding a polypeptide as defined herein. Hybridization probes may be prepared using the CHL2 sequences provided herein to screen cDNA,

genomic or synthetic DNA libraries for related sequences. Regions of the DNA and/or amino acid sequence of CHL2 polypeptide that exhibit significant identity to known sequences are readily determined using sequence alignment algorithms as described herein and those regions may be used to design probes for screening.

The term "highly stringent conditions" refers to those conditions that are designed to permit hybridization of DNA strands whose sequences are highly complementary, and to exclude hybridization of significantly mismatched DNAs. Hybridization stringency is principally determined by temperature, ionic strength, and the concentration of denaturing agents such as formamide. Examples of "highly stringent conditions" for hybridization and washing are 0.015 M sodium CHL2oride, 0.0015 M sodium citrate at 65-68°C or 0.015 M sodium CHL2oride, 0.0015 M sodium citrate, and 50% formamide at 42°C. *See* Sambrook, Fritsch & Maniatis, *Molecular Cloning: A Laboratory Manual* (2nd ed., Cold Spring Harbor Laboratory, 1989); Anderson *et al.*, *Nucleic Acid Hybridisation: A Practical Approach* Ch. 4 (IRL Press Limited).

More stringent conditions (such as higher temperature, lower ionic strength, higher formamide, or other denaturing agent) may also be used – however, the rate of hybridization will be affected. Other agents may be included in the hybridization and washing buffers for the purpose of reducing non-specific and/or background hybridization. Examples are 0.1% bovine serum albumin, 0.1% polyvinyl-pyrrolidone, 0.1% sodium pyrophosphate, 0.1% sodium dodecylsulfate, $NaDodSO_4$, (SDS), ficoll, Denhardt's solution, sonicated salmon sperm DNA (or another non-complementary DNA), and dextran sulfate, although other suitable agents can also be used. The concentration and types of these additives can be changed without substantially affecting the stringency of the hybridization conditions. Hybridization experiments are usually carried out at pH 6.8-7.4; however, at typical ionic strength conditions, the rate of hybridization is nearly independent of pH. *See* Anderson *et al.*, *Nucleic Acid Hybridisation: A Practical Approach* Ch. 4 (IRL Press Limited).

Factors affecting the stability of DNA duplex include base composition, length, and degree of base pair mismatch. Hybridization conditions can be

adjusted by one skilled in the art in order to accommodate these variables and allow DNAs of different sequence relatedness to form hybrids. The melting temperature of a perfectly matched DNA duplex can be estimated by the following equation:

$T_m(°C) = 81.5 + 16.6(\log[Na+]) + 0.41(\%G+C) - 600/N - 0.72(\%formamide)$

where N is the length of the duplex formed, [Na+] is the molar concentration of the sodium ion in the hybridization or washing solution, %G+C is the percentage of (guanine+cytosine) bases in the hybrid. For imperfectly matched hybrids, the melting temperature is reduced by approximately 1°C for each 1% mismatch.

The term "moderately stringent conditions" refers to conditions under which a DNA duplex with a greater degree of base pair mismatching than could occur under "highly stringent conditions" is able to form. Examples of typical "moderately stringent conditions" are 0.015 M sodium CHL2oride, 0.0015 M sodium citrate at 50-65°C or 0.015 M sodium CHL2oride, 0.0015 M sodium citrate, and 20% formamide at 37-50°C. By way of example, "moderately stringent conditions" of 50°C in 0.015 M sodium ion will allow about a 21% mismatch.

It will be appreciated by those skilled in the art that there is no absolute distinction between "highly stringent conditions" and "moderately stringent conditions." For example, at 0.015 M sodium ion (no formamide), the melting temperature of perfectly matched long DNA is about 71°C. With a wash at 65°C (at the same ionic strength), this would allow for approximately a 6% mismatch. To capture more distantly related sequences, one skilled in the art can simply lower the temperature or raise the ionic strength.

A good estimate of the melting temperature in 1M NaCl* for oligonucleotide probes up to about 20nt is given by:

Tm = 2°C per A-T base pair + 4°C per G-C base pair

*The sodium ion concentration in 6X salt sodium citrate (SSC) is 1M. *See* Suggs *et al., Developmental Biology Using Purified Genes* 683 (Brown and Fox, eds., 1981).

High stringency washing conditions for oligonucleotides are usually at a temperature of 0-5°C below the Tm of the oligonucleotide in 6X SSC, 0.1% SDS.

In another embodiment, related nucleic acid molecules comprise or consist of a nucleotide sequence that is at least about 70 percent identical to the nucleotide sequence as shown in either SEQ ID NO: 1 or SEQ ID NO: 4, or comprise or consist essentially of a nucleotide sequence encoding a polypeptide that is at least about 70 percent identical to the polypeptide as set forth in either SEQ ID NO: 2 or SEQ ID NO: 5. In preferred embodiments, the nucleotide sequences are about 75 percent, or about 80 percent, or about 85 percent, or about 90 percent, or about 95, 96, 97, 98, or 99 percent identical to the nucleotide sequence as shown in either SEQ ID NO: 1 or SEQ ID NO: 4, or the nucleotide sequences encode a polypeptide that is about 75 percent, or about 80 percent, or about 85 percent, or about 90 percent, or about 95, 96, 97, 98, or 99 percent identical to the polypeptide sequence as set forth in either SEQ ID NO: 2 or SEQ ID NO: 5. Related nucleic acid molecules encode polypeptides possessing at least one activity of the polypeptide set forth in either SEQ ID NO: 2 or SEQ ID NO: 5.

Differences in the nucleic acid sequence may result in conservative and/or non-conservative modifications of the amino acid sequence relative to the amino acid sequence of any of SEQ ID NO: 2 or SEQ ID NO: 5.

Conservative modifications to the amino acid sequence of any of SEQ ID NO: 2 or SEQ ID NO: 5 (and the corresponding modifications to the encoding nucleotides) will produce a polypeptide having functional and chemical characteristics similar to those of CHL2 polypeptides. In contrast, substantial modifications in the functional and/or chemical characteristics of CHL2 polypeptides may be accomplished by selecting substitutions in the amino acid sequence of any of SEQ ID NO: 2 or SEQ ID NO: 5 that differ significantly in their effect on maintaining (a) the structure of the molecular backbone in the area of the substitution, for example, as a sheet or helical conformation, (b) the charge or hydrophobicity of the molecule at the target site, or (c) the bulk of the side chain.

For example, a "conservative amino acid substitution" may involve a substitution of a native amino acid residue with a nonnative residue such that there is little or no effect on the polarity or charge of the amino acid residue at that position. Furthermore, any native residue in the polypeptide may also be substituted with alanine, as has been previously described for "alanine scanning mutagenesis."

Conservative amino acid substitutions also encompass non-naturally occurring amino acid residues that are typically incorporated by chemical peptide synthesis rather than by synthesis in biological systems. These include peptidomimetics, and other reversed or inverted forms of amino acid moieties.

Naturally occurring residues may be divided into classes based on common side chain properties:

1) hydrophobic: norleucine, Met, Ala, Val, Leu, Ile;

2) neutral hydrophilic: Cys, Ser, Thr;

3) acidic: Asp, Glu;

4) basic: Asn, Gln, His, Lys, Arg;

5) residues that influence chain orientation: Gly, Pro; and

6) aromatic: Trp, Tyr, Phe.

For example, non-conservative substitutions may involve the exchange of a member of one of these classes for a member from another class. Such substituted residues may be introduced into regions of the human CHL2 polypeptide that are homologous with non-human CHL2 polypeptides, or into the non-homologous regions of the molecule.

In making such changes, the hydropathic index of amino acids may be considered. Each amino acid has been assigned a hydropathic index on the basis of its hydrophobicity and charge characteristics. The hydropathic indices are: isoleucine (+4.5); valine (+4.2); leucine (+3.8); phenylalanine (+2.8); cysteine/cystine (+2.5); methionine (+1.9); alanine (+1.8); glycine (-0.4); threonine (-0.7); serine (-0.8); tryptophan (-0.9); tyrosine (-1.3); proline (-1.6); histidine (-3.2); glutamate (-3.5); glutamine (-3.5); aspartate (-3.5); asparagine (-3.5); lysine (-3.9); and arginine (-4.5).

The importance of the hydropathic amino acid index in conferring interactive biological function on a protein is generally understood in the art (Kyte *et al.*, 1982, *J. Mol. Biol.* 157:105-31). It is known that certain amino acids may be substituted for other amino acids having a similar hydropathic index or score and still retain a similar biological activity. In making changes based upon the hydropathic index, the substitution of amino acids whose hydropathic indices are within ±2 is preferred, those which are within ±1 are particularly preferred, and those within ±0.5 are even more particularly preferred.

It is also understood in the art that the substitution of like amino acids can be made effectively on the basis of hydrophilicity, particularly where the biologically functionally equivalent protein or peptide thereby created is intended for use in immunological embodiments, as in the present case. The greatest local average hydrophilicity of a protein, as governed by the hydrophilicity of its adjacent amino acids, correlates with its immunogenicity and antigenicity, *i.e.*, with a biological property of the protein.

The following hydrophilicity values have been assigned to these amino acid residues: arginine (+3.0); lysine (+3.0); aspartate (+3.0 ± 1); glutamate (+3.0 ± 1); serine (+0.3); asparagine (+0.2); glutamine (+0.2); glycine (0); threonine (-0.4); proline (-0.5 ± 1); alanine (-0.5); histidine (-0.5); cysteine (-1.0); methionine (-1.3); valine (-1.5); leucine (-1.8); isoleucine (-1.8); tyrosine (-2.3); phenylalanine (-2.5); and tryptophan (-3.4). In making changes based upon similar hydrophilicity values, the substitution of amino acids whose hydrophilicity values are within ±2 is preferred, those which are within ±1 are particularly preferred, and those within ±0.5 are even more particularly preferred. One may also identify epitopes from primary amino acid sequences on the basis of hydrophilicity. These regions are also referred to as "epitopic core regions."

Desired amino acid substitutions (whether conservative or non-conservative) can be determined by those skilled in the art at the time such substitutions are desired. For example, amino acid substitutions can be used to identify important residues of the CHL2 polypeptide, or to increase or decrease

the affinity of the CHL2 polypeptides described herein. Exemplary amino acid substitutions are set forth in Table I.

Table I

Amino Acid Substitutions

| Original Residues | Exemplary Substitutions | Preferred Substitutions |
|---|---|---|
| Ala | Val, Leu, Ile | Val |
| Arg | Lys, Gln, Asn | Lys |
| Asn | Gln | Gln |
| Asp | Glu | Glu |
| Cys | Ser, Ala | Ser |
| Gln | Asn | Asn |
| Glu | Asp | Asp |
| Gly | Pro, Ala | Ala |
| His | Asn, Gln, Lys, Arg | Arg |
| Ile | Leu, Val, Met, Ala, Phe, Norleucine | Leu |
| Leu | Norleucine, Ile, Val, Met, Ala, Phe | Ile |
| Lys | Arg, 1,4 Diamino-butyric Acid, Gln, Asn | Arg |
| Met | Leu, Phe, Ile | Leu |
| Phe | Leu, Val, Ile, Ala, Tyr | Leu |
| Pro | Ala | Gly |
| Ser | Thr, Ala, Cys | Thr |
| Thr | Ser | Ser |
| Trp | Tyr, Phe | Tyr |
| Tyr | Trp, Phe, Thr, Ser | Phe |

| | | |
|---|---|---|
| Val | Ile, Met, Leu, Phe, Ala, Norleucine | Leu |

A skilled artisan will be able to determine suitable variants of the polypeptide as set forth in either SEQ ID NO: 2 or SEQ ID NO: 5 using well-known techniques. For identifying suitable areas of the molecule that may be changed without destroying biological activity, one skilled in the art may target areas not believed to be important for activity. For example, when similar polypeptides with similar activities from the same species or from other species are known, one skilled in the art may compare the amino acid sequence of a CHL2 polypeptide to such similar polypeptides. With such a comparison, one can identify residues and portions of the molecules that are conserved among similar polypeptides. It will be appreciated that changes in areas of the CHL2 molecule that are not conserved relative to such similar polypeptides would be less likely to adversely affect the biological activity and/or structure of a CHL2 polypeptide. One skilled in the art would also know that, even in relatively conserved regions, one may substitute chemically similar amino acids for the naturally occurring residues while retaining activity (conservative amino acid residue substitutions). Therefore, even areas that may be important for biological activity or for structure may be subject to conservative amino acid substitutions without destroying the biological activity or without adversely affecting the polypeptide structure.

Additionally, one skilled in the art can review structure-function studies identifying residues in similar polypeptides that are important for activity or structure. In view of such a comparison, one can predict the importance of amino acid residues in a CHL2 polypeptide that correspond to amino acid residues that are important for activity or structure in similar polypeptides. One skilled in the art may opt for chemically similar amino acid substitutions for such predicted important amino acid residues of CHL2 polypeptides.

One skilled in the art can also analyze the three-dimensional structure and amino acid sequence in relation to that structure in similar polypeptides. In view of such information, one skilled in the art may predict the alignment of amino acid

residues of CHL2 polypeptide with respect to its three dimensional structure. One skilled in the art may choose not to make radical changes to amino acid residues predicted to be on the surface of the protein, since such residues may be involved in important interactions with other molecules. Moreover, one skilled in the art may generate test variants containing a single amino acid substitution at each amino acid residue. The variants could be screened using activity assays known to those with skill in the art. Such variants could be used to gather information about suitable variants. For example, if one discovered that a change to a particular amino acid residue resulted in destroyed, undesirably reduced, or unsuitable activity, variants with such a change would be avoided. In other words, based on information gathered from such routine experiments, one skilled in the art can readily determine the amino acids where further substitutions should be avoided either alone or in combination with other mutations.

A number of scientific publications have been devoted to the prediction of secondary structure. *See* Moult, 1996, *Curr. Opin. Biotechnol.* 7:422-27; Chou *et al.*, 1974, *Biochemistry* 13:222-45; Chou *et al.*, 1974, *Biochemistry* 113:211-22; Chou *et al.*, 1978, *Adv. Enzymol. Relat. Areas Mol. Biol.* 47:45-48; Chou *et al.*, 1978, *Ann. Rev. Biochem.* 47:251-276; and Chou *et al.*, 1979, *Biophys. J.* 26:367-84. Moreover, computer programs are currently available to assist with predicting secondary structure. One method of predicting secondary structure is based upon homology modeling. For example, two polypeptides or proteins which have a sequence identity of greater than 30%, or similarity greater than 40%, often have similar structural topologies. The recent growth of the protein structural database (PDB) has provided enhanced predictability of secondary structure, including the potential number of folds within the structure of a polypeptide or protein. *See* Holm *et al.*, 1999, *Nucleic Acids Res.* 27:244-47. It has been suggested that there are a limited number of folds in a given polypeptide or protein and that once a critical number of structures have been resolved, structural prediction will become dramatically more accurate (Brenner *et al.*, 1997, *Curr. Opin. Struct. Biol.* 7:369-76).

Additional methods of predicting secondary structure include "threading" (Jones, 1997, *Curr. Opin. Struct. Biol.* 7:377-87; Sippl *et al.*, 1996, *Structure* 4:15-19), "profile analysis" (Bowie *et al.*, 1991, *Science*, 253:164-70; Gribskov *et al.*, 1990, *Methods Enzymol.* 183:146-59; Gribskov *et al.*, 1987, *Proc. Nat. Acad. Sci. U.S.A.* 84:4355-58), and "evolutionary linkage" (*See* Holm *et al.*, *supra*, and Brenner *et al.*, *supra*).

Preferred CHL2 polypeptide variants include glycosylation variants wherein the number and/or type of glycosylation sites have been altered compared to the amino acid sequence set forth in either SEQ ID NO: 2 or SEQ ID NO: 5. In one embodiment, CHL2 polypeptide variants comprise a greater or a lesser number of N-linked glycosylation sites than the amino acid sequence set forth in either SEQ ID NO: 2 or SEQ ID NO: 5. An N-linked glycosylation site is characterized by the sequence: Asn-X-Ser or Asn-X-Thr, wherein the amino acid residue designated as X may be any amino acid residue except proline. The substitution of amino acid residues to create this sequence provides a potential new site for the addition of an N-linked carbohydrate chain. Alternatively, substitutions that eliminate this sequence will remove an existing N-linked carbohydrate chain. Also provided is a rearrangement of N-linked carbohydrate chains wherein one or more N-linked glycosylation sites (typically those that are naturally occurring) are eliminated and one or more new N-linked sites are created. Additional preferred CHL2 variants include cysteine variants, wherein one or more cysteine residues are deleted or substituted with another amino acid (*e.g.*, serine) as compared to the amino acid sequence set forth in either SEQ ID NO: 2 or SEQ ID NO: 5. Cysteine variants are useful when CHL2 polypeptides must be refolded into a biologically active conformation such as after the isolation of insoluble inclusion bodies. Cysteine variants generally have fewer cysteine residues than the native protein, and typically have an even number to minimize interactions resulting from unpaired cysteines.

In other embodiments, related nucleic acid molecules comprise or consist of a nucleotide sequence encoding a polypeptide as set forth in either SEQ ID NO: 2 or SEQ ID NO: 5 with at least one amino acid insertion and wherein the

polypeptide has an activity of the polypeptide set forth in either SEQ ID NO: 2 or SEQ ID NO: 5, or a nucleotide sequence encoding a polypeptide as set forth in either SEQ ID NO: 2 or SEQ ID NO: 5 with at least one amino acid deletion and wherein the polypeptide has an activity of the polypeptide set forth in either SEQ ID NO: 2 or SEQ ID NO: 5. Related nucleic acid molecules also comprise or consist of a nucleotide sequence encoding a polypeptide as set forth in either SEQ ID NO: 2 or SEQ ID NO: 5 wherein the polypeptide has a carboxyl- and/or amino-terminal truncation and further wherein the polypeptide has an activity of the polypeptide set forth in either SEQ ID NO: 2 or SEQ ID NO: 5. Related nucleic acid molecules also comprise or consist of a nucleotide sequence encoding a polypeptide as set forth in either SEQ ID NO: 2 or SEQ ID NO: 5 with at least one modification selected from the group consisting of amino acid substitutions, amino acid insertions, amino acid deletions, carboxyl-terminal truncations, and amino-terminal truncations and wherein the polypeptide has an activity of the polypeptide set forth in either SEQ ID NO: 2 or SEQ ID NO: 5.

In addition, the polypeptide comprising the amino acid sequence of any of SEQ ID NO: 2 or SEQ ID NO: 5, or other CHL2 polypeptide, may be fused to a homologous polypeptide to form a homodimer or to a heterologous polypeptide to form a heterodimer. Heterologous peptides and polypeptides include, but are not limited to: an epitope to allow for the detection and/or isolation of a CHL2 fusion polypeptide; a transmembrane receptor protein or a portion thereof, such as an extracellular domain or a transmembrane and intracellular domain; a ligand or a portion thereof which binds to a transmembrane receptor protein; an enzyme or portion thereof which is catalytically active; a polypeptide or peptide which promotes oligomerization, such as a leucine zipper domain; a polypeptide or peptide which increases stability, such as an immunoglobulin constant region; and a polypeptide which has a therapeutic activity different from the polypeptide comprising the amino acid sequence as set forth in either SEQ ID NO: 2 or SEQ ID NO: 5, or other CHL2 polypeptide.

Fusions can be made either at the amino-terminus or at the carboxyl-terminus of the polypeptide comprising the amino acid sequence set forth in either

SEQ ID NO: 2 or SEQ ID NO: 5, or other CHL2 polypeptide. Fusions may be direct with no linker or adapter molecule or may be through a linker or adapter molecule. A linker or adapter molecule may be one or more amino acid residues, typically from about 20 to about 50 amino acid residues. A linker or adapter molecule may also be designed with a cleavage site for a DNA restriction endonuclease or for a protease to allow for the separation of the fused moieties. It will be appreciated that once constructed, the fusion polypeptides can be derivatized according to the methods described herein.

In a further embodiment of the invention, the polypeptide comprising the amino acid sequence of any of SEQ ID NO: 2 OR SEQ ID NO: 5, or other CHL2 polypeptide, is fused to one or more domains of an Fc region of human IgG. Antibodies comprise two functionally independent parts, a variable domain known as "Fab," that binds an antigen, and a constant domain known as "Fc," that is involved in effector functions such as complement activation and attack by phagocytic cells. An Fc has a long serum half-life, whereas an Fab is short-lived. Capon *et al.*, 1989, *Nature* 337:525-31. When constructed together with a therapeutic protein, an Fc domain can provide longer half-life or incorporate such functions as Fc receptor binding, protein A binding, complement fixation, and perhaps even placental transfer. *Id.* Table II summarizes the use of certain Fc fusions known in the art.

Table II

Fc Fusion with Therapeutic Proteins

| Form of Fc | Fusion partner | Therapeutic implications | Reference |
|---|---|---|---|
| IgG1 | N-terminus of CD30-L | Hodgkin's disease; anaplastic lymphoma; T-cell leukemia | U.S. Patent No. 5,480,981 |
| Murine Fcγ2a | IL-10 | anti-inflammatory; transplant rejection | Zheng *et al.*, 199<br>*Immunol.* 154:55 |
| IgG1 | TNF receptor | septic shock | Fisher *et al.*, 199<br>*Engl. J. Med.* 334<br>1702; Van Zee *et*<br>1996, *J. Immunol.*<br>156:2221-30 |

| | | | |
|---|---|---|---|
| IgG, IgA, IgM, or IgE (excluding the first domain) | TNF receptor | inflammation, autoimmune disorders | U.S. Patent No. 5,808,029 |
| IgG1 | CD4 receptor | AIDS | Capon *et al.*, 198 *Nature* 337: 525- |
| IgG1, IgG3 | N-terminus of IL-2 | anti-cancer, antiviral | Harvill *et al.*, 199 *Immunotech.* 1:9 |
| IgG1 | C-terminus of OPG | osteoarthritis; bone density | WO 97/23614 |
| IgG1 | N-terminus of leptin | anti-obesity | PCT/US 97/2318 December 11, 19 |
| Human Ig Cγ1 | CTLA-4 | autoimmune disorders | Linsley, 1991, *J. Med.*, 174:561-69 |

In one example, a human IgG hinge, CH2, and CH3 region may be fused at either the amino-terminus or carboxyl-terminus of the CHL2 polypeptides using methods known to the skilled artisan. In another example, a human IgG hinge, CH2, and CH3 region may be fused at either the amino-terminus or carboxyl-terminus of a CHL2 polypeptide fragment (*e.g.*, the predicted extracellular portion of CHL2 polypeptide).

The resulting CHL2 fusion polypeptide may be purified by use of a Protein A affinity column. Peptides and proteins fused to an Fc region have been found to exhibit a substantially greater half-life *in vivo* than the unfused counterpart. Also, a fusion to an Fc region allows for dimerization/multimerization of the fusion polypeptide. The Fc region may be a naturally occurring Fc region, or may be altered to improve certain qualities, such as therapeutic qualities, circulation time, or reduced aggregation.

Identity and similarity of related nucleic acid molecules and polypeptides are readily calculated by known methods. Such methods include, but are not limited to those described in *Computational Molecular Biology* (A.M. Lesk, ed., Oxford University Press 1988); *Biocomputing: Informatics and Genome Projects* (D.W. Smith, ed., Academic Press 1993); *Computer Analysis of Sequence Data* (Part 1, A.M. Griffin and H.G. Griffin, eds., Humana Press 1994); G. von Heinle, *Sequence Analysis in Molecular Biology* (Academic Press 1987); *Sequence*

*Analysis Primer* (M. Gribskov and J. Devereux, eds., M. Stockton Press 1991); and Carillo *et al.*, 1988, *SIAM J. Applied Math.*, 48:1073.

Preferred methods to determine identity and/or similarity are designed to give the largest match between the sequences tested. Methods to determine identity and similarity are described in publicly available computer programs. Preferred computer program methods to determine identity and similarity between two sequences include, but are not limited to, the GCG program package, including GAP (Devereux *et al.*, 1984, *Nucleic Acids Res.* 12:387; Genetics Computer Group, University of Wisconsin, Madison, WI), BLASTP, BLASTN, and FASTA (Altschul *et al.*, 1990, *J. Mol. Biol.* 215:403-10). The BLASTX program is publicly available from the National Center for Biotechnology Information (NCBI) and other sources (Altschul *et al.*, *BLAST Manual* (NCB NLM NIH, Bethesda, MD); Altschul *et al.*, 1990, *supra*). The well-known Smith Waterman algorithm may also be used to determine identity.

Certain alignment schemes for aligning two amino acid sequences may result in the matching of only a short region of the two sequences, and this small aligned region may have very high sequence identity even though there is no significant relationship between the two full-length sequences. Accordingly, in a preferred embodiment, the selected alignment method (GAP program) will result in an alignment that spans at least 50 contiguous amino acids of the claimed polypeptide.

For example, using the computer algorithm GAP (Genetics Computer Group, University of Wisconsin, Madison, WI), two polypeptides for which the percent sequence identity is to be determined are aligned for optimal matching of their respective amino acids (the "matched span," as determined by the algorithm). A gap opening penalty (which is calculated as 3X the average diagonal; the "average diagonal" is the average of the diagonal of the comparison matrix being used; the "diagonal" is the score or number assigned to each perfect amino acid match by the particular comparison matrix) and a gap extension penalty (which is usually 0.1X the gap opening penalty), as well as a comparison matrix such as PAM 250 or BLOSUM 62 are used in conjunction with the

algorithm. A standard comparison matrix is also used by the algorithm (*see* Dayhoff *et al.*, 5 *Atlas of Protein Sequence and Structure* (Supp. 3 1978)(PAM250 comparison matrix); Henikoff *et al.*, 1992, *Proc. Natl. Acad. Sci USA* 89:10915-19 (BLOSUM 62 comparison matrix)).

Preferred parameters for polypeptide sequence comparison include the following:

Algorithm: Needleman and Wunsch, 1970, *J. Mol. Biol.* 48:443-53;
Comparison matrix: BLOSUM 62 (Henikoff *et al.*, *supra*);
Gap Penalty: 12
Gap Length Penalty: 4
Threshold of Similarity: 0

The GAP program is useful with the above parameters. The aforementioned parameters are the default parameters for polypeptide comparisons (along with no penalty for end gaps) using the GAP algorithm.

Preferred parameters for nucleic acid molecule sequence comparison include the following:

Algorithm: Needleman and Wunsch, *supra*;
Comparison matrix: matches = +10, mismatch = 0
Gap Penalty: 50
Gap Length Penalty: 3

The GAP program is also useful with the above parameters. The aforementioned parameters are the default parameters for nucleic acid molecule comparisons.

Other exemplary algorithms, gap opening penalties, gap extension penalties, comparison matrices, and thresholds of similarity may be used, including those set forth in the Program Manual, Wisconsin Package, Version 9, September, 1997. The particular choices to be made will be apparent to those of skill in the art and will depend on the specific comparison to be made, such as

DNA-to-DNA, protein-to-protein, protein-to-DNA; and additionally, whether the comparison is between given pairs of sequences (in which case GAP or BestFit are generally preferred) or between one sequence and a large database of sequences (in which case FASTA or BLASTA are preferred).

Nucleic Acid Molecules

The nucleic acid molecules encoding a polypeptide comprising the amino acid sequence of a CHL2 polypeptide can readily be obtained in a variety of ways including, without limitation, chemical synthesis, cDNA or genomic library screening, expression library screening, and/or PCR amplification of cDNA.

Recombinant DNA methods used herein are generally those set forth in Sambrook *et al.*, *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Laboratory Press, 1989) and/or *Current Protocols in Molecular Biology* (Ausubel *et al.*, eds., Green Publishers Inc. and Wiley and Sons 1994). The invention provides for nucleic acid molecules as described herein and methods for obtaining such molecules.

Where a gene encoding the amino acid sequence of a CHL2 polypeptide has been identified from one species, all or a portion of that gene may be used as a probe to identify orthologs or related genes from the same species. The probes or primers may be used to screen cDNA libraries from various tissue sources believed to express the CHL2 polypeptide. In addition, part or all of a nucleic acid molecule having the sequence as set forth in either SEQ ID NO: 1 or SEQ ID NO: 4 may be used to screen a genomic library to identify and isolate a gene encoding the amino acid sequence of a CHL2 polypeptide. Typically, conditions of moderate or high stringency will be employed for screening to minimize the number of false positives obtained from the screening.

Nucleic acid molecules encoding the amino acid sequence of CHL2 polypeptides may also be identified by expression cloning which employs the detection of positive clones based upon a property of the expressed protein. Typically, nucleic acid libraries are screened by the binding an antibody or other binding partner (*e.g.*, receptor or ligand) to cloned proteins that are expressed and

displayed on a host cell surface. The antibody or binding partner is modified with a detectable label to identify those cells expressing the desired clone.

Recombinant expression techniques conducted in accordance with the descriptions set forth below may be followed to produce these polynucleotides and to express the encoded polypeptides. For example, by inserting a nucleic acid sequence that encodes the amino acid sequence of a CHL2 polypeptide into an appropriate vector, one skilled in the art can readily produce large quantities of the desired nucleotide sequence. The sequences can then be used to generate detection probes or amplification primers. Alternatively, a polynucleotide encoding the amino acid sequence of a CHL2 polypeptide can be inserted into an expression vector. By introducing the expression vector into an appropriate host, the encoded CHL2 polypeptide may be produced in large amounts.

Another method for obtaining a suitable nucleic acid sequence is the polymerase chain reaction (PCR). In this method, cDNA is prepared from poly(A)+RNA or total RNA using the enzyme reverse transcriptase. Two primers, typically complementary to two separate regions of cDNA encoding the amino acid sequence of a CHL2 polypeptide, are then added to the cDNA along with a polymerase such as *Taq* polymerase, and the polymerase amplifies the cDNA region between the two primers.

Another means of preparing a nucleic acid molecule encoding the amino acid sequence of a CHL2 polypeptide is chemical synthesis using methods well known to the skilled artisan such as those described by Engels *et al.*, 1989, *Angew. Chem. Intl. Ed.* 28:716-34. These methods include, *inter alia*, the phosphotriester, phosphoramidite, and H-phosphonate methods for nucleic acid synthesis. A preferred method for such chemical synthesis is polymer-supported synthesis using standard phosphoramidite chemistry. Typically, the DNA encoding the amino acid sequence of a CHL2 polypeptide will be several hundred nucleotides in length. Nucleic acids larger than about 100 nucleotides can be synthesized as several fragments using these methods. The fragments can then be ligated together to form the full-length nucleotide sequence of a CHL2 gene. Usually, the DNA fragment encoding the amino-terminus of the polypeptide will

have an ATG, which encodes a methionine residue. This methionine may or may not be present on the mature form of the CHL2 polypeptide, depending on whether the polypeptide produced in the host cell is designed to be secreted from that cell. Other methods known to the skilled artisan may be used as well.

In certain embodiments, nucleic acid variants contain codons which have been altered for optimal expression of a CHL2 polypeptide in a given host cell. Particular codon alterations will depend upon the CHL2 polypeptide and host cell selected for expression. Such "codon optimization" can be carried out by a variety of methods, for example, by selecting codons which are preferred for use in highly expressed genes in a given host cell. Computer algorithms which incorporate codon frequency tables such as "Eco_high.Cod" for codon preference of highly expressed bacterial genes may be used and are provided by the University of Wisconsin Package Version 9.0 (Genetics Computer Group, Madison, WI). Other useful codon frequency tables include "Celegans_high.cod," "Celegans_low.cod," "Drosophila_high.cod," "Human_high.cod," "Maize_high.cod," and "Yeast_high.cod."

In some cases, it may be desirable to prepare nucleic acid molecules encoding CHL2 polypeptide variants. Nucleic acid molecules encoding variants may be produced using site directed mutagenesis, PCR amplification, or other appropriate methods, where the primer(s) have the desired point mutations (*see* Sambrook *et al.*, *supra*, and Ausubel *et al.*, *supra*, for descriptions of mutagenesis techniques). Chemical synthesis using methods described by Engels *et al.*, *supra*, may also be used to prepare such variants. Other methods known to the skilled artisan may be used as well.

Vectors and Host Cells

A nucleic acid molecule encoding the amino acid sequence of a CHL2 polypeptide is inserted into an appropriate expression vector using standard ligation techniques. The vector is typically selected to be functional in the particular host cell employed (*i.e.*, the vector is compatible with the host cell machinery such that amplification of the gene and/or expression of the gene can

occur). A nucleic acid molecule encoding the amino acid sequence of a CHL2 polypeptide may be amplified/expressed in prokaryotic, yeast, insect (baculovirus systems) and/or eukaryotic host cells. Selection of the host cell will depend in part on whether a CHL2 polypeptide is to be post-translationally modified (*e.g.*, glycosylated and/or phosphorylated). If so, yeast, insect, or mammalian host cells are preferable. For a review of expression vectors, *see Meth. Enz.*, vol. 185 (D.V. Goeddel, ed., Academic Press 1990).

Typically, expression vectors used in any of the host cells will contain sequences for plasmid maintenance and for cloning and expression of exogenous nucleotide sequences. Such sequences, collectively referred to as "flanking sequences" in certain embodiments will typically include one or more of the following nucleotide sequences: a promoter, one or more enhancer sequences, an origin of replication, a transcriptional termination sequence, a complete intron sequence containing a donor and acceptor splice site, a sequence encoding a leader sequence for polypeptide secretion, a ribosome binding site, a polyadenylation sequence, a polylinker region for inserting the nucleic acid encoding the polypeptide to be expressed, and a selectable marker element. Each of these sequences is discussed below.

Optionally, the vector may contain a "tag"-encoding sequence, *i.e.*, an oligonucleotide molecule located at the 5' or 3' end of the CHL2 polypeptide coding sequence; the oligonucleotide sequence encodes polyHis (such as hexaHis), or another "tag" such as FLAG, HA (hemaglutinin influenza virus), or *myc* for which commercially available antibodies exist. This tag is typically fused to the polypeptide upon expression of the polypeptide, and can serve as a means for affinity purification of the CHL2 polypeptide from the host cell. Affinity purification can be accomplished, for example, by column chromatography using antibodies against the tag as an affinity matrix. Optionally, the tag can subsequently be removed from the purified CHL2 polypeptide by various means such as using certain peptidases for cleavage.

Flanking sequences may be homologous (*i.e.*, from the same species and/or strain as the host cell), heterologous (*i.e.*, from a species other than the host

cell species or strain), hybrid (*i.e.*, a combination of flanking sequences from more than one source), or synthetic, or the flanking sequences may be native sequences which normally function to regulate CHL2 polypeptide expression. As such, the source of a flanking sequence may be any prokaryotic or eukaryotic organism, any vertebrate or invertebrate organism, or any plant, provided that the flanking sequence is functional in, and can be activated by, the host cell machinery.

Flanking sequences useful in the vectors of this invention may be obtained by any of several methods well known in the art. Typically, flanking sequences useful herein – other than the CHL2 gene flanking sequences – will have been previously identified by mapping and/or by restriction endonuclease digestion and can thus be isolated from the proper tissue source using the appropriate restriction endonucleases. In some cases, the full nucleotide sequence of a flanking sequence may be known. Here, the flanking sequence may be synthesized using the methods described herein for nucleic acid synthesis or cloning.

Where all or only a portion of the flanking sequence is known, it may be obtained using PCR and/or by screening a genomic library with a suitable oligonucleotide and/or flanking sequence fragment from the same or another species. Where the flanking sequence is not known, a fragment of DNA containing a flanking sequence may be isolated from a larger piece of DNA that may contain, for example, a coding sequence or even another gene or genes. Isolation may be accomplished by restriction endonuclease digestion to produce the proper DNA fragment followed by isolation using agarose gel purification, Qiagen® column chromatography (Chatsworth, CA), or other methods known to the skilled artisan. The selection of suitable enzymes to accomplish this purpose will be readily apparent to one of ordinary skill in the art.

An origin of replication is typically a part of those prokaryotic expression vectors purchased commercially, and the origin aids in the amplification of the vector in a host cell. Amplification of the vector to a certain copy number can, in some cases, be important for the optimal expression of a CHL2 polypeptide. If the vector of choice does not contain an origin of replication site, one may be

chemically synthesized based on a known sequence, and ligated into the vector. For example, the origin of replication from the plasmid pBR322 (New England Biolabs, Beverly, MA) is suitable for most gram-negative bacteria and various origins (*e.g.*, SV40, polyoma, adenovirus, vesicular stomatitus virus (VSV), or papillomaviruses such as HPV or BPV) are useful for cloning vectors in mammalian cells. Generally, the origin of replication component is not needed for mammalian expression vectors (for example, the SV40 origin is often used only because it contains the early promoter).

A transcription termination sequence is typically located 3' of the end of a polypeptide coding region and serves to terminate transcription. Usually, a transcription termination sequence in prokaryotic cells is a G-C rich fragment followed by a poly-T sequence. While the sequence is easily cloned from a library or even purchased commercially as part of a vector, it can also be readily synthesized using methods for nucleic acid synthesis such as those described herein.

A selectable marker gene element encodes a protein necessary for the survival and growth of a host cell grown in a selective culture medium. Typical selection marker genes encode proteins that (a) confer resistance to antibiotics or other toxins, *e.g.*, ampicillin, tetracycline, or kanamycin for prokaryotic host cells; (b) complement auxotrophic deficiencies of the cell; or (c) supply critical nutrients not available from complex media. Preferred selectable markers are the kanamycin resistance gene, the ampicillin resistance gene, and the tetracycline resistance gene. A neomycin resistance gene may also be used for selection in prokaryotic and eukaryotic host cells.

Other selection genes may be used to amplify the gene that will be expressed. Amplification is the process wherein genes that are in greater demand for the production of a protein critical for growth are reiterated in tandem within the chromosomes of successive generations of recombinant cells. Examples of suitable selectable markers for mammalian cells include dihydrofolate reductase (DHFR) and thymidine kinase. The mammalian cell transformants are placed under selection pressure wherein only the transformants are uniquely adapted to

survive by virtue of the selection gene present in the vector. Selection pressure is imposed by culturing the transformed cells under conditions in which the concentration of selection agent in the medium is successively changed, thereby leading to the amplification of both the selection gene and the DNA that encodes a CHL2 polypeptide. As a result, increased quantities of CHL2 polypeptide are synthesized from the amplified DNA.

A ribosome binding site is usually necessary for translation initiation of mRNA and is characterized by a Shine-Dalgarno sequence (prokaryotes) or a Kozak sequence (eukaryotes). The element is typically located 3' to the promoter and 5' to the coding sequence of a CHL2 polypeptide to be expressed. The Shine-Dalgarno sequence is varied but is typically a polypurine (*i.e.*, having a high A-G content). Many Shine-Dalgarno sequences have been identified, each of which can be readily synthesized using methods set forth herein and used in a prokaryotic vector.

A leader, or signal, sequence may be used to direct a CHL2 polypeptide out of the host cell. Typically, a nucleotide sequence encoding the signal sequence is positioned in the coding region of a CHL2 nucleic acid molecule, or directly at the 5' end of a CHL2 polypeptide coding region. Many signal sequences have been identified, and any of those that are functional in the selected host cell may be used in conjunction with a CHL2 nucleic acid molecule. Therefore, a signal sequence may be homologous (naturally occurring) or heterologous to the CHL2 nucleic acid molecule. Additionally, a signal sequence may be chemically synthesized using methods described herein. In most cases, the secretion of a CHL2 polypeptide from the host cell via the presence of a signal peptide will result in the removal of the signal peptide from the secreted CHL2 polypeptide. The signal sequence may be a component of the vector, or it may be a part of a CHL2 nucleic acid molecule that is inserted into the vector.

Included within the scope of this invention is the use of either a nucleotide sequence encoding a native CHL2 polypeptide signal sequence joined to a CHL2 polypeptide coding region or a nucleotide sequence encoding a heterologous signal sequence joined to a CHL2 polypeptide coding region. The heterologous

signal sequence selected should be one that is recognized and processed, *i.e.*, cleaved by a signal peptidase, by the host cell. For prokaryotic host cells that do not recognize and process the native CHL2 polypeptide signal sequence, the signal sequence is substituted by a prokaryotic signal sequence selected, for example, from the group of the alkaline phosphatase, penicillinase, or heat-stable enterotoxin II leaders. For yeast secretion, the native CHL2 polypeptide signal sequence may be substituted by the yeast invertase, alpha factor, or acid phosphatase leaders. In mammalian cell expression the native signal sequence is satisfactory, although other mammalian signal sequences may be suitable.

In some cases, such as where glycosylation is desired in a eukaryotic host cell expression system, one may manipulate the various presequences to improve glycosylation or yield. For example, one may alter the peptidase cleavage site of a particular signal peptide, or add pro-sequences, which also may affect glycosylation. The final protein product may have, in the -1 position (relative to the first amino acid of the mature protein) one or more additional amino acids incident to expression, which may not have been totally removed. For example, the final protein product may have one or two amino acid residues found in the peptidase cleavage site, attached to the amino-terminus. Alternatively, use of some enzyme cleavage sites may result in a slightly truncated form of the desired CHL2 polypeptide, if the enzyme cuts at such area within the mature polypeptide.

In many cases, transcription of a nucleic acid molecule is increased by the presence of one or more introns in the vector; this is particularly true where a polypeptide is produced in eukaryotic host cells, especially mammalian host cells. The introns used may be naturally occurring within the CHL2 gene especially where the gene used is a full-length genomic sequence or a fragment thereof. Where the intron is not naturally occurring within the gene (as for most cDNAs), the intron may be obtained from another source. The position of the intron with respect to flanking sequences and the CHL2 gene is generally important, as the intron must be transcribed to be effective. Thus, when a CHL2 cDNA molecule is being transcribed, the preferred position for the intron is 3' to the transcription start site and 5' to the poly-A transcription termination sequence. Preferably, the

intron or introns will be located on one side or the other (*i.e.*, 5' or 3') of the cDNA such that it does not interrupt the coding sequence. Any intron from any source, including viral, prokaryotic and eukaryotic (plant or animal) organisms, may be used to practice this invention, provided that it is compatible with the host cell into which it is inserted. Also included herein are synthetic introns. Optionally, more than one intron may be used in the vector.

The expression and cloning vectors of the present invention will typically contain a promoter that is recognized by the host organism and operably linked to the molecule encoding the CHL2 polypeptide. Promoters are untranscribed sequences located upstream (*i.e.*, 5') to the start codon of a structural gene (generally within about 100 to 1000 bp) that control the transcription of the structural gene. Promoters are conventionally grouped into one of two classes: inducible promoters and constitutive promoters. Inducible promoters initiate increased levels of transcription from DNA under their control in response to some change in culture conditions, such as the presence or absence of a nutrient or a change in temperature. Constitutive promoters, on the other hand, initiate continual gene product production; that is, there is little or no control over gene expression. A large number of promoters, recognized by a variety of potential host cells, are well known. A suitable promoter is operably linked to the DNA encoding CHL2 polypeptide by removing the promoter from the source DNA by restriction enzyme digestion and inserting the desired promoter sequence into the vector. The native CHL2 promoter sequence may be used to direct amplification and/or expression of a CHL2 nucleic acid molecule. A heterologous promoter is preferred, however, if it permits greater transcription and higher yields of the expressed protein as compared to the native promoter, and if it is compatible with the host cell system that has been selected for use.

Promoters suitable for use with prokaryotic hosts include the beta-lactamase and lactose promoter systems; alkaline phosphatase; a tryptophan (trp) promoter system; and hybrid promoters such as the tac promoter. Other known bacterial promoters are also suitable. Their sequences have been published,

thereby enabling one skilled in the art to ligate them to the desired DNA sequence, using linkers or adapters as needed to supply any useful restriction sites.

Suitable promoters for use with yeast hosts are also well known in the art. Yeast enhancers are advantageously used with yeast promoters. Suitable promoters for use with mammalian host cells are well known and include, but are not limited to, those obtained from the genomes of viruses such as polyoma virus, fowlpox virus, adenovirus (such as Adenovirus 2), bovine papilloma virus, avian sarcoma virus, cytomegalovirus, retroviruses, hepatitis-B virus and most preferably Simian Virus 40 (SV40). Other suitable mammalian promoters include heterologous mammalian promoters, for example, heat-shock promoters and the actin promoter.

Additional promoters which may be of interest in controlling CHL2 gene expression include, but are not limited to: the SV40 early promoter region (Bernoist and Chambon, 1981, *Nature* 290:304-10); the CMV promoter; the promoter contained in the 3' long terminal repeat of Rous sarcoma virus (Yamamoto, *et al.*, 1980, *Cell* 22:787-97); the herpes thymidine kinase promoter (Wagner *et al.*, 1981, *Proc. Natl. Acad. Sci. U.S.A.* 78:1444-45); the regulatory sequences of the metallothionine gene (Brinster *et al.*, 1982, *Nature* 296:39-42); prokaryotic expression vectors such as the beta-lactamase promoter (Villa-Kamaroff *et al.*, 1978, *Proc. Natl. Acad. Sci. U.S.A.*, 75:3727-31); or the tac promoter (DeBoer *et al.*, 1983, *Proc. Natl. Acad. Sci. U.S.A.*, 80:21-25). Also of interest are the following animal transcriptional control regions, which exhibit tissue specificity and have been utilized in transgenic animals: the elastase I gene control region which is active in pancreatic acinar cells (Swift *et al.*, 1984, *Cell* 38:639-46; Ornitz *et al.*, 1986, *Cold Spring Harbor Symp. Quant. Biol.* 50:399-409 (1986); MacDonald, 1987, *Hepatology* 7:425-515); the insulin gene control region which is active in pancreatic beta cells (Hanahan, 1985, *Nature* 315:115-22); the immunoglobulin gene control region which is active in lymphoid cells (Grosschedl *et al.*, 1984, *Cell* 38:647-58; Adames *et al.*, 1985, *Nature* 318:533-38; Alexander *et al.*, 1987, *Mol. Cell. Biol.*, 7:1436-44); the mouse mammary tumor virus control region which is active in testicular, breast, lymphoid and mast cells

(Leder *et al.*, 1986, *Cell* 45:485-95); the albumin gene control region which is active in liver (Pinkert *et al.*, 1987, *Genes and Devel.* 1:268-76); the alpha-feto-protein gene control region which is active in liver (Krumlauf *et al.*, 1985, *Mol. Cell. Biol.*, 5:1639-48; Hammer *et al.*, 1987, *Science* 235:53-58); the alpha 1-antitrypsin gene control region which is active in the liver (Kelsey *et al.*, 1987, *Genes and Devel.* 1:161-71); the beta-globin gene control region which is active in myeloid cells (Mogram *et al.*, 1985, *Nature* 315:338-40; Kollias *et al.*, 1986, *Cell* 46:89-94); the myelin basic protein gene control region which is active in oligodendrocyte cells in the brain (Readhead *et al.*, 1987, *Cell* 48:703-12); the myosin light chain-2 gene control region which is active in skeletal muscle (Sani, 1985, *Nature* 314:283-86); and the gonadotropic releasing hormone gene control region which is active in the hypothalamus (Mason *et al.*, 1986, *Science* 234:1372-78).

An enhancer sequence may be inserted into the vector to increase the transcription of a DNA encoding a CHL2 polypeptide of the present invention by higher eukaryotes. Enhancers are cis-acting elements of DNA, usually about 10-300 bp in length, that act on the promoter to increase transcription. Enhancers are relatively orientation and position independent. They have been found 5' and 3' to the transcription unit. Several enhancer sequences available from mammalian genes are known (*e.g.*, globin, elastase, albumin, alpha-feto-protein and insulin). Typically, however, an enhancer from a virus will be used. The SV40 enhancer, the cytomegalovirus early promoter enhancer, the polyoma enhancer, and adenovirus enhancers are exemplary enhancing elements for the activation of eukaryotic promoters. While an enhancer may be spliced into the vector at a position 5' or 3' to a CHL2 nucleic acid molecule, it is typically located at a site 5' from the promoter.

Expression vectors of the invention may be constructed from a starting vector such as a commercially available vector. Such vectors may or may not contain all of the desired flanking sequences. Where one or more of the flanking sequences described herein are not already present in the vector, they may be

individually obtained and ligated into the vector. Methods used for obtaining each of the flanking sequences are well known to one skilled in the art.

Preferred vectors for practicing this invention are those which are compatible with bacterial, insect, and mammalian host cells. Such vectors include, *inter alia*, pCRII, pCR3, and pcDNA3.1 (Invitrogen, San Diego, CA), pBSII (Stratagene, La Jolla, CA), pET15 (Novagen, Madison, WI), pGEX (Pharmacia Biotech, Piscataway, NJ), pEGFP-N2 (Clontech, Palo Alto, CA), pETL (BlueBacII, Invitrogen), pDSR-alpha (PCT Pub. No. WO 90/14363) and pFastBacDual (Gibco-BRL, Grand Island, NY).

Additional suitable vectors include, but are not limited to, cosmids, plasmids, or modified viruses, but it will be appreciated that the vector system must be compatible with the selected host cell. Such vectors include, but are not limited to plasmids such as Bluescript® plasmid derivatives (a high copy number ColE1-based phagemid, Stratagene Cloning Systems, La Jolla CA), PCR cloning plasmids designed for cloning Taq-amplified PCR products (*e.g.*, TOPO™ TA Cloning® Kit, PCR2.1® plasmid derivatives, Invitrogen, Carlsbad, CA), and mammalian, yeast or virus vectors such as a baculovirus expression system (pBacPAK plasmid derivatives, Clontech, Palo Alto, CA).

After the vector has been constructed and a nucleic acid molecule encoding a CHL2 polypeptide has been inserted into the proper site of the vector, the completed vector may be inserted into a suitable host cell for amplification and/or polypeptide expression. The transformation of an expression vector for a CHL2 polypeptide into a selected host cell may be accomplished by well known methods including methods such as transfection, infection, calcium CHL2oride, electroporation, microinjection, lipofection, DEAE-dextran method, or other known techniques. The method selected will in part be a function of the type of host cell to be used. These methods and other suitable methods are well known to the skilled artisan, and are set forth, for example, in Sambrook *et al.*, *supra*.

Host cells may be prokaryotic host cells (such as *E. coli*) or eukaryotic host cells (such as a yeast, insect, or vertebrate cell). The host cell, when cultured under appropriate conditions, synthesizes a CHL2 polypeptide which can

subsequently be collected from the culture medium (if the host cell secretes it into the medium) or directly from the host cell producing it (if it is not secreted). The selection of an appropriate host cell will depend upon various factors, such as desired expression levels, polypeptide modifications that are desirable or necessary for activity (such as glycosylation or phosphorylation) and ease of folding into a biologically active molecule.

A number of suitable host cells are known in the art and many are available from the American Type Culture Collection (ATCC), Manassas, VA. Examples include, but are not limited to, mammalian cells, such as Chinese hamster ovary cells (CHO), CHO DHFR(-) cells (Urlaub *et al.*, 1980, *Proc. Natl. Acad. Sci. U.S.A.* 97:4216-20), human embryonic kidney (HEK) 293 or 293T cells, or 3T3 cells. The selection of suitable mammalian host cells and methods for transformation, culture, amplification, screening, product production, and purification are known in the art. Other suitable mammalian cell lines, are the monkey COS-1 and COS-7 cell lines, and the CV-1 cell line. Further exemplary mammalian host cells include primate cell lines and rodent cell lines, including transformed cell lines. Normal diploid cells, cell strains derived from *in vitro* culture of primary tissue, as well as primary explants, are also suitable. Candidate cells may be genotypically deficient in the selection gene, or may contain a dominantly acting selection gene. Other suitable mammalian cell lines include but are not limited to, mouse neuroblastoma N2A cells, HeLa, mouse L-929 cells, 3T3 lines derived from Swiss, Balb-c or NIH mice, BHK or HaK hamster cell lines. Each of these cell lines is known by and available to those skilled in the art of protein expression.

Similarly useful as host cells suitable for the present invention are bacterial cells. For example, the various strains of *E. coli* (*e.g.*, HB101, DH5α, DH10, and MC1061) are well-known as host cells in the field of biotechnology. Various strains of *B. subtilis*, *Pseudomonas spp.*, other *Bacillus spp.*, *Streptomyces spp.*, and the like may also be employed in this method.

Many strains of yeast cells known to those skilled in the art are also available as host cells for the expression of the polypeptides of the present

invention. Preferred yeast cells include, for example, *Saccharomyces cerivisae* and *Pichia pastoris*.

Additionally, where desired, insect cell systems may be utilized in the methods of the present invention. Such systems are described, for example, in Kitts *et al.*, 1993, *Biotechniques*, 14:810-17; Lucklow, 1993, *Curr. Opin. Biotechnol.* 4:564-72; and Lucklow *et al.*, 1993, *J. Virol.*, 67:4566-79. Preferred insect cells are Sf-9 and Hi5 (Invitrogen).

One may also use transgenic animals to express glycosylated CHL2 polypeptides. For example, one may use a transgenic milk-producing animal (a cow or goat, for example) and obtain the present glycosylated polypeptide in the animal milk. One may also use plants to produce CHL2 polypeptides, however, in general, the glycosylation occurring in plants is different from that produced in mammalian cells, and may result in a glycosylated product which is not suitable for human therapeutic use.

Polypeptide Production

Host cells comprising a CHL2 polypeptide expression vector may be cultured using standard media well known to the skilled artisan. The media will usually contain all nutrients necessary for the growth and survival of the cells. Suitable media for culturing *E. coli* cells include, for example, Luria Broth (LB) and/or Terrific Broth (TB). Suitable media for culturing eukaryotic cells include Roswell Park Memorial Institute medium 1640 (RPMI 1640), Minimal Essential Medium (MEM) and/or Dulbecco's Modified Eagle Medium (DMEM), all of which may be supplemented with serum and/or growth factors as necessary for the particular cell line being cultured. A suitable medium for insect cultures is Grace's medium supplemented with yeastolate, lactalbumin hydrolysate, and/or fetal calf serum as necessary.

Typically, an antibiotic or other compound useful for selective growth of transfected or transformed cells is added as a supplement to the media. The compound to be used will be dictated by the selectable marker element present on the plasmid with which the host cell was transformed. For example, where the

selectable marker element is kanamycin resistance, the compound added to the culture medium will be kanamycin. Other compounds for selective growth include ampicillin, tetracycline, and neomycin.

The amount of a CHL2 polypeptide produced by a host cell can be evaluated using standard methods known in the art. Such methods include, without limitation, Western blot analysis, SDS-polyacrylamide gel electrophoresis, non-denaturing gel electrophoresis, High Performance Liquid Chromatography (HPLC) separation, immunoprecipitation, and/or activity assays such as DNA binding gel shift assays.

If a CHL2 polypeptide has been designed to be secreted from the host cells, the majority of polypeptide may be found in the cell culture medium. If however, the CHL2 polypeptide is not secreted from the host cells, it will be present in the cytoplasm and/or the nucleus (for eukaryotic host cells) or in the cytosol (for gram-negative bacteria host cells).

For a CHL2 polypeptide situated in the host cell cytoplasm and/or nucleus (for eukaryotic host cells) or in the cytosol (for bacterial host cells), the intracellular material (including inclusion bodies for gram-negative bacteria) can be extracted from the host cell using any standard technique known to the skilled artisan. For example, the host cells can be lysed to release the contents of the periplasm/cytoplasm by French press, homogenization, and/or sonication followed by centrifugation.

If a CHL2 polypeptide has formed inclusion bodies in the cytosol, the inclusion bodies can often bind to the inner and/or outer cellular membranes and thus will be found primarily in the pellet material after centrifugation. The pellet material can then be treated at pH extremes or with a chaotropic agent such as a detergent, guanidine, guanidine derivatives, urea, or urea derivatives in the presence of a reducing agent such as dithiothreitol at alkaline pH or tris carboxyethyl phosphine at acid pH to release, break apart, and solubilize the inclusion bodies. The solubilized CHL2 polypeptide can then be analyzed using gel electrophoresis, immunoprecipitation, or the like. If it is desired to isolate the

CHL2 polypeptide, isolation may be accomplished using standard methods such as those described herein and in Marston *et al.*, 1990, *Meth. Enz.*, 182:264-75.

In some cases, a CHL2 polypeptide may not be biologically active upon isolation. Various methods for "refolding" or converting the polypeptide to its tertiary structure and generating disulfide linkages can be used to restore biological activity. Such methods include exposing the solubilized polypeptide to a pH usually above 7 and in the presence of a particular concentration of a chaotrope. The selection of chaotrope is very similar to the choices used for inclusion body solubilization, but usually the chaotrope is used at a lower concentration and is not necessarily the same as chaotropes used for the solubilization. In most cases the refolding/oxidation solution will also contain a reducing agent or the reducing agent plus its oxidized form in a specific ratio to generate a particular redox potential allowing for disulfide shuffling to occur in the formation of the protein's cysteine bridges. Some of the commonly used redox couples include cysteine/cystamine, glutathione (GSH)/dithiobis GSH, cupric CHL2oride, dithiothreitol(DTT)/dithiane DTT, and 2-2-mercaptoethanol(bME)/dithio-b(ME). In many instances, a cosolvent may be used or may be needed to increase the efficiency of the refolding, and the more common reagents used for this purpose include glycerol, polyethylene glycol of various molecular weights, arginine and the like.

If inclusion bodies are not formed to a significant degree upon expression of a CHL2 polypeptide, then the polypeptide will be found primarily in the supernatant after centrifugation of the cell homogenate. The polypeptide may be further isolated from the supernatant using methods such as those described herein.

The purification of a CHL2 polypeptide from solution can be accomplished using a variety of techniques. If the polypeptide has been synthesized such that it contains a tag such as Hexahistidine (CHL2 polypeptide/hexaHis) or other small peptide such as FLAG (Eastman Kodak Co., New Haven, CT) or *myc* (Invitrogen, Carlsbad, CA) at either its carboxyl- or amino-terminus, it may be purified in a one-step process by passing the solution

through an affinity column where the column matrix has a high affinity for the tag.

For example, polyhistidine binds with great affinity and specificity to nickel. Thus, an affinity column of nickel (such as the Qiagen® nickel columns) can be used for purification of CHL2 polypeptide/polyHis. *See, e.g., Current Protocols in Molecular Biology* § 10.11.8 (Ausubel *et al.*, eds., Green Publishers Inc. and Wiley and Sons 1993).

Additionally, CHL2 polypeptides may be purified through the use of a monoclonal antibody that is capable of specifically recognizing and binding to a CHL2 polypeptide.

Other suitable procedures for purification include, without limitation, affinity chromatography, immunoaffinity chromatography, ion exchange chromatography, molecular sieve chromatography, HPLC, electrophoresis (including native gel electrophoresis) followed by gel elution, and preparative isoelectric focusing ("Isoprime" machine/technique, Hoefer Scientific, San Francisco, CA). In some cases, two or more purification techniques may be combined to achieve increased purity.

CHL2 polypeptides may also be prepared by chemical synthesis methods (such as solid phase peptide synthesis) using techniques known in the art such as those set forth by Merrifield *et al.*, 1963, *J. Am. Chem. Soc.* 85:2149; Houghten *et al.*, 1985, *Proc Natl Acad. Sci. USA* 82:5132; and Stewart and Young, *Solid Phase Peptide Synthesis* (Pierce Chemical Co. 1984). Such polypeptides may be synthesized with or without a methionine on the amino-terminus. Chemically synthesized CHL2 polypeptides may be oxidized using methods set forth in these references to form disulfide bridges. Chemically synthesized CHL2 polypeptides are expected to have comparable biological activity to the corresponding CHL2 polypeptides produced recombinantly or purified from natural sources, and thus may be used interchangeably with a recombinant or natural CHL2 polypeptide.

Another means of obtaining CHL2 polypeptide is via purification from biological samples such as source tissues and/or fluids in which the CHL2 polypeptide is naturally found. Such purification can be conducted using methods

for protein purification as described herein. The presence of the CHL2 polypeptide during purification may be monitored, for example, using an antibody prepared against recombinantly produced CHL2 polypeptide or peptide fragments thereof.

A number of additional methods for producing nucleic acids and polypeptides are known in the art, and the methods can be used to produce polypeptides having specificity for CHL2 polypeptide. *See, e.g.*, Roberts *et al.*, 1997, *Proc. Natl. Acad. Sci. U.S.A.* 94:12297-303, which describes the production of fusion proteins between an mRNA and its encoded peptide. *See also,* Roberts, 1999, *Curr. Opin. Chem. Biol.* 3:268-73. Additionally, U.S. Patent No. 5,824,469 describes methods for obtaining oligonucleotides capable of carrying out a specific biological function. The procedure involves generating a heterogeneous pool of oligonucleotides, each having a 5' randomized sequence, a central preselected sequence, and a 3' randomized sequence. The resulting heterogeneous pool is introduced into a population of cells that do not exhibit the desired biological function. Subpopulations of the cells are then screened for those that exhibit a predetermined biological function. From that subpopulation, oligonucleotides capable of carrying out the desired biological function are isolated.

U.S. Patent Nos. 5,763,192; 5,814,476; 5,723,323; and 5,817,483 describe processes for producing peptides or polypeptides. This is done by producing stochastic genes or fragments thereof, and then introducing these genes into host cells which produce one or more proteins encoded by the stochastic genes. The host cells are then screened to identify those clones producing peptides or polypeptides having the desired activity.

Another method for producing peptides or polypeptides is described in PCT/US98/20094 (WO99/15650) filed by Athersys, Inc. Known as "Random Activation of Gene Expression for Gene Discovery" (RAGE-GD), the process involves the activation of endogenous gene expression or over-expression of a gene by *in situ* recombination methods. For example, expression of an endogenous gene is activated or increased by integrating a regulatory sequence

into the target cell which is capable of activating expression of the gene by non-homologous or illegitimate recombination. The target DNA is first subjected to radiation, and a genetic promoter inserted. The promoter eventually locates a break at the front of a gene, initiating transcription of the gene. This results in expression of the desired peptide or polypeptide.

It will be appreciated that these methods can also be used to create comprehensive CHL2 polypeptide expression libraries, which can subsequently be used for high throughput phenotypic screening in a variety of assays, such as biochemical assays, cellular assays, and whole organism assays (*e.g.*, plant, mouse, etc.).

Synthesis

It will be appreciated by those skilled in the art that the nucleic acid and polypeptide molecules described herein may be produced by recombinant and other means.

Selective Binding Agents

The term "selective binding agent" refers to a molecule that has specificity for one or more CHL2 polypeptides. Suitable selective binding agents include, but are not limited to, antibodies and derivatives thereof, polypeptides, and small molecules. Suitable selective binding agents may be prepared using methods known in the art. An exemplary CHL2 polypeptide selective binding agent of the present invention is capable of binding a certain portion of the CHL2 polypeptide thereby inhibiting the binding of the polypeptide to a CHL2 polypeptide receptor.

Selective binding agents such as antibodies and antibody fragments that bind CHL2 polypeptides are within the scope of the present invention. The antibodies may be polyclonal including monospecific polyclonal; monoclonal (MAbs); recombinant; chimeric; humanized, such as CDR-grafted; human; single chain; and/or bispecific; as well as fragments; variants; or derivatives thereof. Antibody fragments include those portions of the antibody that bind to an epitope on the CHL2 polypeptide. Examples of such fragments include Fab and F(ab')

fragments generated by enzymatic cleavage of full-length antibodies. Other binding fragments include those generated by recombinant DNA techniques, such as the expression of recombinant plasmids containing nucleic acid sequences encoding antibody variable regions.

Polyclonal antibodies directed toward a CHL2 polypeptide generally are produced in animals (*e.g.*, rabbits or mice) by means of multiple subcutaneous or intraperitoneal injections of CHL2 polypeptide and an adjuvant. It may be useful to conjugate a CHL2 polypeptide to a carrier protein that is immunogenic in the species to be immunized, such as keyhole limpet hemocyanin, serum, albumin, bovine thyroglobulin, or soybean trypsin inhibitor. Also, aggregating agents such as alum are used to enhance the immune response. After immunization, the animals are bled and the serum is assayed for anti-CHL2 antibody titer.

Monoclonal antibodies directed toward CHL2 polypeptides are produced using any method that provides for the production of antibody molecules by continuous cell lines in culture. Examples of suitable methods for preparing monoclonal antibodies include the hybridoma methods of Kohler *et al.*, 1975, *Nature* 256:495-97 and the human B-cell hybridoma method (Kozbor, 1984, *J. Immunol.* 133:3001; Brodeur *et al.*, *Monoclonal Antibody Production Techniques and Applications* 51-63 (Marcel Dekker, Inc., 1987). Also provided by the invention are hybridoma cell lines that produce monoclonal antibodies reactive with CHL2 polypeptides.

Monoclonal antibodies of the invention may be modified for use as therapeutics. One embodiment is a "chimeric" antibody in which a portion of the heavy (H) and/or light (L) chain is identical with or homologous to a corresponding sequence in antibodies derived from a particular species or belonging to a particular antibody class or subclass, while the remainder of the chain(s) is/are identical with or homologous to a corresponding sequence in antibodies derived from another species or belonging to another antibody class or subclass. Also included are fragments of such antibodies, so long as they exhibit the desired biological activity. *See* U.S. Patent No. 4,816,567; Morrison *et al.*, 1985, *Proc. Natl. Acad. Sci.* 81:6851-55.

In another embodiment, a monoclonal antibody of the invention is a "humanized" antibody. Methods for humanizing non-human antibodies are well known in the art. *See* U.S. Patent Nos. 5,585,089 and 5,693,762. Generally, a humanized antibody has one or more amino acid residues introduced into it from a source that is non-human. Humanization can be performed, for example, using methods described in the art (Jones *et al.*, 1986, *Nature* 321:522-25; Riechmann *et al.*, 1998, *Nature* 332:323-27; Verhoeyen *et al.*, 1988, *Science* 239:1534-36), by substituting at least a portion of a rodent complementarity-determining region (CDR) for the corresponding regions of a human antibody.

Also encompassed by the invention are human antibodies that bind CHL2 polypeptides. Using transgenic animals (*e.g.*, mice) that are capable of producing a repertoire of human antibodies in the absence of endogenous immunoglobulin production such antibodies are produced by immunization with a CHL2 polypeptide antigen (*i.e.*, having at least 6 contiguous amino acids), optionally conjugated to a carrier. *See, e.g.*, Jakobovits *et al.*, 1993, *Proc. Natl. Acad. Sci.* 90:2551-55; Jakobovits *et al.*, 1993, *Nature* 362:255-58; Bruggermann *et al.*, 1993, *Year in Immuno.* 7:33. In one method, such transgenic animals are produced by incapacitating the endogenous loci encoding the heavy and light immunoglobulin chains therein, and inserting loci encoding human heavy and light chain proteins into the genome thereof. Partially modified animals, that is those having less than the full complement of modifications, are then cross-bred to obtain an animal having all of the desired immune system modifications. When administered an immunogen, these transgenic animals produce antibodies with human (rather than, *e.g.*, murine) amino acid sequences, including variable regions which are immunospecific for these antigens. *See* PCT App. Nos. PCT/US96/05928 and PCT/US93/06926. Additional methods are described in U.S. Patent No. 5,545,807, PCT App. Nos. PCT/US91/245 and PCT/GB89/01207, and in European Patent Nos. 546073B1 and 546073A1. Human antibodies can also be produced by the expression of recombinant DNA in host cells or by expression in hybridoma cells as described herein.

In an alternative embodiment, human antibodies can also be produced from phage-display libraries (Hoogenboom *et al.*, 1991, *J. Mol. Biol.* 227:381; Marks *et al.*, 1991, *J. Mol. Biol.* 222:581). These processes mimic immune selection through the display of antibody repertoires on the surface of filamentous bacteriophage, and subsequent selection of phage by their binding to an antigen of choice. One such technique is described in PCT App. No. PCT/US98/17364, which describes the isolation of high affinity and functional agonistic antibodies for MPL- and msk- receptors using such an approach.

Chimeric, CDR grafted, and humanized antibodies are typically produced by recombinant methods. Nucleic acids encoding the antibodies are introduced into host cells and expressed using materials and procedures described herein. In a preferred embodiment, the antibodies are produced in mammalian host cells, such as CHO cells. Monoclonal (*e.g.*, human) antibodies may be produced by the expression of recombinant DNA in host cells or by expression in hybridoma cells as described herein.

The anti-CHL2 antibodies of the invention may be employed in any known assay method, such as competitive binding assays, direct and indirect sandwich assays, and immunoprecipitation assays (Sola, *Monoclonal Antibodies: A Manual of Techniques* 147-158 (CRC Press, Inc., 1987)) for the detection and quantitation of CHL2 polypeptides. The antibodies will bind CHL2 polypeptides with an affinity that is appropriate for the assay method being employed.

For diagnostic applications, in certain embodiments, anti-CHL2 antibodies may be labeled with a detectable moiety. The detectable moiety can be any one that is capable of producing, either directly or indirectly, a detectable signal. For example, the detectable moiety may be a radioisotope, such as $^{3}H$, $^{14}C$, $^{32}P$, $^{35}S$, $^{125}I$, $^{99}Tc$, $^{111}In$, or $^{67}Ga$; a fluorescent or chemiluminescent compound, such as fluorescein isothiocyanate, rhodamine, or luciferin; or an enzyme, such as alkaline phosphatase, β-galactosidase, or horseradish peroxidase (Bayer, *et al.*, 1990, *Meth. Enz.* 184:138-63).

Competitive binding assays rely on the ability of a labeled standard (*e.g.*, a CHL2 polypeptide, or an immunologically reactive portion thereof) to compete

with the test sample analyte (an CHL2 polypeptide) for binding with a limited amount of anti-CHL2 antibody. The amount of a CHL2 polypeptide in the test sample is inversely proportional to the amount of standard that becomes bound to the antibodies. To facilitate determining the amount of standard that becomes bound, the antibodies typically are insolubilized before or after the competition, so that the standard and analyte that are bound to the antibodies may conveniently be separated from the standard and analyte which remain unbound.

Sandwich assays typically involve the use of two antibodies, each capable of binding to a different immunogenic portion, or epitope, of the protein to be detected and/or quantitated. In a sandwich assay, the test sample analyte is typically bound by a first antibody which is immobilized on a solid support, and thereafter a second antibody binds to the analyte, thus forming an insoluble three-part complex. *See, e.g.*, U.S. Patent No. 4,376,110. The second antibody may itself be labeled with a detectable moiety (direct sandwich assays) or may be measured using an anti-immunoglobulin antibody that is labeled with a detectable moiety (indirect sandwich assays). For example, one type of sandwich assay is an enzyme-linked immunosorbent assay (ELISA), in which case the detectable moiety is an enzyme.

The selective binding agents, including anti-CHL2 antibodies, are also useful for *in vivo* imaging. An antibody labeled with a detectable moiety may be administered to an animal, preferably into the bloodstream, and the presence and location of the labeled antibody in the host assayed. The antibody may be labeled with any moiety that is detectable in an animal, whether by nuclear magnetic resonance, radiology, or other detection means known in the art.

Selective binding agents of the invention, including antibodies, may be used as therapeutics. These therapeutic agents are generally agonists or antagonists, in that they either enhance or reduce, respectively, at least one of the biological activities of a CHL2 polypeptide. In one embodiment, antagonist antibodies of the invention are antibodies or binding fragments thereof which are capable of specifically binding to a CHL2 polypeptide and which are capable of inhibiting or eliminating the functional activity of a CHL2 polypeptide *in vivo* or

*in vitro*. In preferred embodiments, the selective binding agent, *e.g.*, an antagonist antibody, will inhibit the functional activity of a CHL2 polypeptide by at least about 50%, and preferably by at least about 80%. In another embodiment, the selective binding agent may be an anti-CHL2 polypeptide antibody that is capable of interacting with a CHL2 polypeptide binding partner (a ligand or receptor) thereby inhibiting or eliminating CHL2 polypeptide activity *in vitro* or *in vivo*. Selective binding agents, including agonist and antagonist anti-CHL2 polypeptide antibodies, are identified by screening assays that are well known in the art.

The invention also relates to a kit comprising CHL2 selective binding agents (such as antibodies) and other reagents useful for detecting CHL2 polypeptide levels in biological samples. Such reagents may include a detectable label, blocking serum, positive and negative control samples, and detection reagents.

Microarrays

It will be appreciated that DNA microarray technology can be utilized in accordance with the present invention. DNA microarrays are miniature, high-density arrays of nucleic acids positioned on a solid support, such as glass. Each cell or element within the array contains numerous copies of a single nucleic acid species that acts as a target for hybridization with a complementary nucleic acid sequence (*e.g.*, mRNA). In expression profiling using DNA microarray technology, mRNA is first extracted from a cell or tissue sample and then converted enzymatically to fluorescently labeled cDNA. This material is hybridized to the microarray and unbound cDNA is removed by washing. The expression of discrete genes represented on the array is then visualized by quantitating the amount of labeled cDNA that is specifically bound to each target nucleic acid molecule. In this way, the expression of thousands of genes can be quantitated in a high throughput, parallel manner from a single sample of biological material.

This high throughput expression profiling has a broad range of applications with respect to the CHL2 molecules of the invention, including, but

not limited to: the identification and validation of CHL2 disease-related genes as targets for therapeutics; molecular toxicology of related CHL2 molecules and inhibitors thereof; stratification of populations and generation of surrogate markers for clinical trials; and enhancing related CHL2 polypeptide small molecule drug discovery by aiding in the identification of selective compounds in high throughput screens.

Chemical Derivatives

Chemically modified derivatives of CHL2 polypeptides may be prepared by one skilled in the art, given the disclosures described herein. CHL2 polypeptide derivatives are modified in a manner that is different – either in the type or location of the molecules naturally attached to the polypeptide. Derivatives may include molecules formed by the deletion of one or more naturally-attached chemical groups. The polypeptide comprising the amino acid sequence of any of SEQ ID NO: 2 or SEQ ID NO: 5, or other CHL2 polypeptide, may be modified by the covalent attachment of one or more polymers. For example, the polymer selected is typically water-soluble so that the protein to which it is attached does not precipitate in an aqueous environment, such as a physiological environment. Included within the scope of suitable polymers is a mixture of polymers. Preferably, for therapeutic use of the end-product preparation, the polymer will be pharmaceutically acceptable.

The polymers each may be of any molecular weight and may be branched or unbranched. The polymers each typically have an average molecular weight of between about 2 kDa to about 100 kDa (the term "about" indicating that in preparations of a water-soluble polymer, some molecules will weigh more, some less, than the stated molecular weight). The average molecular weight of each polymer is preferably between about 5 kDa and about 50 kDa, more preferably between about 12 kDa and about 40 kDa and most preferably between about 20 kDa and about 35 kDa.

Suitable water-soluble polymers or mixtures thereof include, but are not limited to, N-linked or O-linked carbohydrates, sugars, phosphates, polyethylene

glycol (PEG) (including the forms of PEG that have been used to derivatize proteins, including mono-($C_1$-$C_{10}$), alkoxy-, or aryloxy-polyethylene glycol), monomethoxy-polyethylene glycol, dextran (such as low molecular weight dextran of, for example, about 6 kD), cellulose, or other carbohydrate based polymers, poly-(N-vinyl pyrrolidone) polyethylene glycol, propylene glycol homopolymers, polypropylene oxide/ethylene oxide co-polymers, polyoxyethylated polyols (*e.g.*, glycerol), and polyvinyl alcohol. Also encompassed by the present invention are bifunctional crosslinking molecules which may be used to prepare covalently attached CHL2 polypeptide multimers.

In general, chemical derivatization may be performed under any suitable condition used to react a protein with an activated polymer molecule. Methods for preparing chemical derivatives of polypeptides will generally comprise the steps of: (a) reacting the polypeptide with the activated polymer molecule (such as a reactive ester or aldehyde derivative of the polymer molecule) under conditions whereby the polypeptide comprising the amino acid sequence of any of SEQ ID NO: 2 or SEQ ID NO: 5, or other CHL2 polypeptide, becomes attached to one or more polymer molecules, and (b) obtaining the reaction products. The optimal reaction conditions will be determined based on known parameters and the desired result. For example, the larger the ratio of polymer molecules to protein, the greater the percentage of attached polymer molecule. In one embodiment, the CHL2 polypeptide derivative may have a single polymer molecule moiety at the amino-terminus. *See, e.g.*, U.S. Patent No. 5,234,784.

The pegylation of a polypeptide may be specifically carried out using any of the pegylation reactions known in the art. Such reactions are described, for example, in the following references: Francis *et al.*, 1992, *Focus on Growth Factors* 3:4-10; European Patent Nos. 0154316 and 0401384; and U.S. Patent No. 4,179,337. For example, pegylation may be carried out via an acylation reaction or an alkylation reaction with a reactive polyethylene glycol molecule (or an analogous reactive water-soluble polymer) as described herein. For the acylation reactions, a selected polymer should have a single reactive ester group. For reductive alkylation, a selected polymer should have a single reactive aldehyde

group. A reactive aldehyde is, for example, polyethylene glycol propionaldehyde, which is water stable, or mono $C_1$-$C_{10}$ alkoxy or aryloxy derivatives thereof (*see* U.S. Patent No. 5,252,714).

In another embodiment, CHL2 polypeptides may be chemically coupled to biotin. The biotin/CHL2 polypeptide molecules are then allowed to bind to avidin, resulting in tetravalent avidin/biotin/CHL2 polypeptide molecules. CHL2 polypeptides may also be covalently coupled to dinitrophenol (DNP) or trinitrophenol (TNP) and the resulting conjugates precipitated with anti-DNP or anti-TNP-IgM to form decameric conjugates with a valency of 10.

Generally, conditions that may be alleviated or modulated by the administration of the present CHL2 polypeptide derivatives include those described herein for CHL2 polypeptides. However, the CHL2 polypeptide derivatives disclosed herein may have additional activities, enhanced or reduced biological activity, or other characteristics, such as increased or decreased half-life, as compared to the non-derivatized molecules.

Genetically Engineered Non-Human Animals

Additionally included within the scope of the present invention are non-human animals such as mice, rats, or other rodents; rabbits, goats, sheep, or other farm animals, in which the genes encoding native CHL2 polypeptide have been disrupted (*i.e.*, "knocked out") such that the level of expression of CHL2 polypeptide is significantly decreased or completely abolished. Such animals may be prepared using techniques and methods such as those described in U.S. Patent No. 5,557,032.

The present invention further includes non-human animals such as mice, rats, or other rodents; rabbits, goats, sheep, or other farm animals, in which either the native form of a CHL2 gene for that animal or a heterologous CHL2 gene is over-expressed by the animal, thereby creating a "transgenic" animal. Such transgenic animals may be prepared using well known methods such as those described in U.S. Patent No 5,489,743 and PCT Pub. No. WO 94/28122.

The present invention further includes non-human animals in which the promoter for one or more of the CHL2 polypeptides of the present invention is either activated or inactivated (*e.g.*, by using homologous recombination methods) to alter the level of expression of one or more of the native CHL2 polypeptides.

These non-human animals may be used for drug candidate screening. In such screening, the impact of a drug candidate on the animal may be measured. For example, drug candidates may decrease or increase the expression of the CHL2 gene. In certain embodiments, the amount of CHL2 polypeptide that is produced may be measured after the exposure of the animal to the drug candidate. Additionally, in certain embodiments, one may detect the actual impact of the drug candidate on the animal. For example, over-expression of a particular gene may result in, or be associated with, a disease or pathological condition. In such cases, one may test a drug candidate's ability to decrease expression of the gene or its ability to prevent or inhibit a pathological condition. In other examples, the production of a particular metabolic product such as a fragment of a polypeptide, may result in, or be associated with, a disease or pathological condition. In such cases, one may test a drug candidate's ability to decrease the production of such a metabolic product or its ability to prevent or inhibit a pathological condition.

Assaying for Other Modulators of CHL2 Polypeptide Activity

In some situations, it may be desirable to identify molecules that are modulators, *i.e.*, agonists or antagonists, of the activity of CHL2 polypeptide. Natural or synthetic molecules that modulate CHL2 polypeptide may be identified using one or more screening assays, such as those described herein. Such molecules may be administered either in an *ex vivo* manner or in an *in vivo* manner by injection, or by oral delivery, implantation device, or the like.

"Test molecule" refers to a molecule that is under evaluation for the ability to modulate (*i.e.*, increase or decrease) the activity of a CHL2 polypeptide. Most commonly, a test molecule will interact directly with a CHL2 polypeptide. However, it is also contemplated that a test molecule may also modulate CHL2 polypeptide activity indirectly, such as by affecting CHL2 gene expression, or by

binding to a CHL2 polypeptide binding partner (*e.g.*, receptor or ligand). In one embodiment, a test molecule will bind to a CHL2 polypeptide with an affinity constant of at least about $10^{-6}$ M, preferably about $10^{-8}$ M, more preferably about $10^{-9}$ M, and even more preferably about $10^{-10}$ M.

Methods for identifying compounds that interact with CHL2 polypeptides are encompassed by the present invention. In certain embodiments, a CHL2 polypeptide is incubated with a test molecule under conditions that permit the interaction of the test molecule with a CHL2 polypeptide, and the extent of the interaction is measured. The test molecule can be screened in a substantially purified form or in a crude mixture.

In certain embodiments, a CHL2 polypeptide agonist or antagonist may be a protein, peptide, carbohydrate, lipid, or small molecular weight molecule that interacts with CHL2 polypeptide to regulate its activity. Molecules which regulate CHL2 polypeptide expression include nucleic acids which are complementary to nucleic acids encoding a CHL2 polypeptide, or are complementary to nucleic acids sequences which direct or control the expression of CHL2 polypeptide, and which act as anti-sense regulators of expression.

Once a test molecule has been identified as interacting with a CHL2 polypeptide, the molecule may be further evaluated for its ability to increase or decrease CHL2 polypeptide activity. The measurement of the interaction of a test molecule with CHL2 polypeptide may be carried out in several formats, including cell-based binding assays, membrane binding assays, solution-phase assays, and immunoassays. In general, a test molecule is incubated with a CHL2 polypeptide for a specified period of time, and CHL2 polypeptide activity is determined by one or more assays for measuring biological activity.

The interaction of test molecules with CHL2 polypeptides may also be assayed directly using polyclonal or monoclonal antibodies in an immunoassay. Alternatively, modified forms of CHL2 polypeptides containing epitope tags as described herein may be used in solution and immunoassays.

In the event that CHL2 polypeptides display biological activity through an interaction with a binding partner (*e.g.*, a receptor or a ligand), a variety of *in vitro*

assays may be used to measure the binding of a CHL2 polypeptide to the corresponding binding partner (such as a selective binding agent, receptor, or ligand). These assays may be used to screen test molecules for their ability to increase or decrease the rate and/or the extent of binding of a CHL2 polypeptide to its binding partner. In one assay, a CHL2 polypeptide is immobilized in the wells of a microtiter plate. Radiolabeled CHL2 polypeptide binding partner (for example, iodinated CHL2 polypeptide binding partner) and a test molecule can then be added either one at a time (in either order) or simultaneously to the wells. After incubation, the wells can be washed and counted for radioactivity, using a scintillation counter, to determine the extent to which the binding partner bound to the CHL2 polypeptide. Typically, a molecule will be tested over a range of concentrations, and a series of control wells lacking one or more elements of the test assays can be used for accuracy in the evaluation of the results. An alternative to this method involves reversing the "positions" of the proteins, *i.e.*, immobilizing CHL2 polypeptide binding partner to the microtiter plate wells, incubating with the test molecule and radiolabeled CHL2 polypeptide, and determining the extent of CHL2 polypeptide binding. *See, e.g., Current Protocols in Molecular Biology*, chap. 18 (Ausubel *et al.*, eds., Green Publishers Inc. and Wiley and Sons 1995).

As an alternative to radiolabeling, a CHL2 polypeptide or its binding partner may be conjugated to biotin, and the presence of biotinylated protein can then be detected using streptavidin linked to an enzyme, such as horse radish peroxidase (HRP) or alkaline phosphatase (AP), which can be detected colorometrically, or by fluorescent tagging of streptavidin. An antibody directed to a CHL2 polypeptide or to a CHL2 polypeptide binding partner, and which is conjugated to biotin, may also be used for purposes of detection following incubation of the complex with enzyme-linked streptavidin linked to AP or HRP.

A CHL2 polypeptide or a CHL2 polypeptide binding partner can also be immobilized by attachment to agarose beads, acrylic beads, or other types of such inert solid phase substrates. The substrate-protein complex can be placed in a solution containing the complementary protein and the test compound. After

incubation, the beads can be precipitated by centrifugation, and the amount of binding between a CHL2 polypeptide and its binding partner can be assessed using the methods described herein. Alternatively, the substrate-protein complex can be immobilized in a column with the test molecule and complementary protein passing through the column. The formation of a complex between a CHL2 polypeptide and its binding partner can then be assessed using any of the techniques described herein (*e.g.*, radiolabelling or antibody binding).

Another *in vitro* assay that is useful for identifying a test molecule which increases or decreases the formation of a complex between a CHL2 polypeptide binding protein and a CHL2 polypeptide binding partner is a surface plasmon resonance detector system such as the BIAcore assay system (Pharmacia, Piscataway, NJ). The BIAcore system is utilized as specified by the manufacturer. This assay essentially involves the covalent binding of either CHL2 polypeptide or a CHL2 polypeptide binding partner to a dextran-coated sensor chip that is located in a detector. The test compound and the other complementary protein can then be injected, either simultaneously or sequentially, into the chamber containing the sensor chip. The amount of complementary protein that binds can be assessed based on the change in molecular mass that is physically associated with the dextran-coated side of the sensor chip, with the change in molecular mass being measured by the detector system.

In some cases, it may be desirable to evaluate two or more test compounds together for their ability to increase or decrease the formation of a complex between a CHL2 polypeptide and a CHL2 polypeptide binding partner. In these cases, the assays set forth herein can be readily modified by adding such additional test compound(s) either simultaneously with, or subsequent to, the first test compound. The remainder of the steps in the assay are as set forth herein.

*In vitro* assays such as those described herein may be used advantageously to screen large numbers of compounds for an effect on the formation of a complex between a CHL2 polypeptide and CHL2 polypeptide binding partner. The assays may be automated to screen compounds generated in phage display, synthetic peptide, and chemical synthesis libraries.

Compounds which increase or decrease the formation of a complex between a CHL2 polypeptide and a CHL2 polypeptide binding partner may also be screened in cell culture using cells and cell lines expressing either CHL2 polypeptide or CHL2 polypeptide binding partner. Cells and cell lines may be obtained from any mammal, but preferably will be from human or other primate, canine, or rodent sources. The binding of a CHL2 polypeptide to cells expressing CHL2 polypeptide binding partner at the surface is evaluated in the presence or absence of test molecules, and the extent of binding may be determined by, for example, flow cytometry using a biotinylated antibody to a CHL2 polypeptide binding partner. Cell culture assays can be used advantageously to further evaluate compounds that score positive in protein binding assays described herein.

Cell cultures can also be used to screen the impact of a drug candidate. For example, drug candidates may decrease or increase the expression of the CHL2 gene. In certain embodiments, the amount of CHL2 polypeptide or a CHL2 polypeptide fragment that is produced may be measured after exposure of the cell culture to the drug candidate. In certain embodiments, one may detect the actual impact of the drug candidate on the cell culture. For example, the over-expression of a particular gene may have a particular impact on the cell culture. In such cases, one may test a drug candidate's ability to increase or decrease the expression of the gene or its ability to prevent or inhibit a particular impact on the cell culture. In other examples, the production of a particular metabolic product such as a fragment of a polypeptide, may result in, or be associated with, a disease or pathological condition. In such cases, one may test a drug candidate's ability to decrease the production of such a metabolic product in a cell culture.

Internalizing Proteins

The *tat* protein sequence (from HIV) can be used to internalize proteins into a cell. *See, e.g.*, Falwell *et al.*, 1994, *Proc. Natl. Acad. Sci. U.S.A.* 91:664-68. For example, an 11 amino acid sequence (Y-G-R-K-K-R-R-Q-R-R-R; SEQ ID NO: 10) of the HIV tat protein (termed the "protein transduction domain," or TAT

PDT) has been described as mediating delivery across the cytoplasmic membrane and the nuclear membrane of a cell. *See* Schwarze *et al.*, 1999, *Science* 285:1569-72; and Nagahara *et al.*, 1998, *Nat. Med.* 4:1449-52. In these procedures, FITC-constructs (FITC-labeled G-G-G-G-Y-G-R-K-K-R-R-Q-R-R-R; SEQ ID NO: 11), which penetrate tissues following intraperitoneal administration, are prepared, and the binding of such constructs to cells is detected by fluorescence-activated cell sorting (FACS) analysis. Cells treated with a *tat*-β-gal fusion protein will demonstrate β-gal activity. Following injection, expression of such a construct can be detected in a number of tissues, including liver, kidney, lung, heart, and brain tissue. It is believed that such constructs undergo some degree of unfolding in order to enter the cell, and as such, may require a refolding following entry into the cell.

It will thus be appreciated that the *tat* protein sequence may be used to internalize a desired polypeptide into a cell. For example, using the *tat* protein sequence, a CHL2 antagonist (such as an anti-CHL2 selective binding agent, small molecule, soluble receptor, or antisense oligonucleotide) can be administered intracellularly to inhibit the activity of a CHL2 molecule. As used herein, the term "CHL2 molecule" refers to both CHL2 nucleic acid molecules and CHL2 polypeptides as defined herein. Where desired, the CHL2 protein itself may also be internally administered to a cell using these procedures. *See also*, Straus, 1999, *Science* 285:1466-67.

Cell Source Identification Using CHL2 Polypeptide

In accordance with certain embodiments of the invention, it may be useful to be able to determine the source of a certain cell type associated with a CHL2 polypeptide. For example, it may be useful to determine the origin of a disease or pathological condition as an aid in selecting an appropriate therapy. In certain embodiments, nucleic acids encoding a CHL2 polypeptide can be used as a probe to identify cells described herein by screening the nucleic acids of the cells with such a probe. In other embodiments, one may use anti-CHL2 polypeptide

antibodies to test for the presence of CHL2 polypeptide in cells, and thus, determine if such cells are of the types described herein.

CHL2 Polypeptide Compositions and Administration

Therapeutic compositions are within the scope of the present invention. Such CHL2 polypeptide pharmaceutical compositions may comprise a therapeutically effective amount of a CHL2 polypeptide or a CHL2 nucleic acid molecule in admixture with a pharmaceutically or physiologically acceptable formulation agent selected for suitability with the mode of administration. Pharmaceutical compositions may comprise a therapeutically effective amount of one or more CHL2 polypeptide selective binding agents in admixture with a pharmaceutically or physiologically acceptable formulation agent selected for suitability with the mode of administration.

Acceptable formulation materials preferably are nontoxic to recipients at the dosages and concentrations employed.

The pharmaceutical composition may contain formulation materials for modifying, maintaining, or preserving, for example, the pH, osmolarity, viscosity, clarity, color, isotonicity, odor, sterility, stability, rate of dissolution or release, adsorption, or penetration of the composition. Suitable formulation materials include, but are not limited to, amino acids (such as glycine, glutamine, asparagine, arginine, or lysine), antimicrobials, antioxidants (such as ascorbic acid, sodium sulfite, or sodium hydrogen-sulfite), buffers (such as borate, bicarbonate, Tris-HCl, citrates, phosphates, or other organic acids), bulking agents (such as mannitol or glycine), chelating agents (such as ethylenediamine tetraacetic acid (EDTA)), complexing agents (such as caffeine, polyvinylpyrrolidone, beta-cyclodextrin, or hydroxypropyl-beta-cyclodextrin), fillers, monosaccharides, disaccharides, and other carbohydrates (such as glucose, mannose, or dextrins), proteins (such as serum albumin, gelatin, or immunoglobulins), coloring, flavoring and diluting agents, emulsifying agents, hydrophilic polymers (such as polyvinylpyrrolidone), low molecular weight polypeptides, salt-forming counterions (such as sodium), preservatives (such as

benzalkonium CHL2oride, benzoic acid, salicylic acid, thimerosal, phenethyl alcohol, methylparaben, propylparaben, CHL2orhexidine, sorbic acid, or hydrogen peroxide), solvents (such as glycerin, propylene glycol, or polyethylene glycol), sugar alcohols (such as mannitol or sorbitol), suspending agents, surfactants or wetting agents (such as pluronics; PEG; sorbitan esters; polysorbates such as polysorbate 20 or polysorbate 80; triton; tromethamine; lecithin; cholesterol or tyloxapal), stability enhancing agents (such as sucrose or sorbitol), tonicity enhancing agents (such as alkali metal halides – preferably sodium or potassium CHL2oride – or mannitol sorbitol), delivery vehicles, diluents, excipients and/or pharmaceutical adjuvants. *See Remington's Pharmaceutical Sciences* (18th Ed., A.R. Gennaro, ed., Mack Publishing Company 1990.

The optimal pharmaceutical composition will be determined by a skilled artisan depending upon, for example, the intended route of administration, delivery format, and desired dosage. *See, e.g., Remington's Pharmaceutical Sciences, supra.* Such compositions may influence the physical state, stability, rate of *in vivo* release, and rate of *in vivo* clearance of the CHL2 molecule.

The primary vehicle or carrier in a pharmaceutical composition may be either aqueous or non-aqueous in nature. For example, a suitable vehicle or carrier for injection may be water, physiological saline solution, or artificial cerebrospinal fluid, possibly supplemented with other materials common in compositions for parenteral administration. Neutral buffered saline or saline mixed with serum albumin are further exemplary vehicles. Other exemplary pharmaceutical compositions comprise Tris buffer of about pH 7.0-8.5, or acetate buffer of about pH 4.0-5.5, which may further include sorbitol or a suitable substitute. In one embodiment of the present invention, CHL2 polypeptide compositions may be prepared for storage by mixing the selected composition having the desired degree of purity with optional formulation agents (*Remington's Pharmaceutical Sciences, supra*) in the form of a lyophilized cake or an aqueous solution. Further, the CHL2 polypeptide product may be formulated as a lyophilizate using appropriate excipients such as sucrose.

The CHL2 polypeptide pharmaceutical compositions can be selected for parenteral delivery. Alternatively, the compositions may be selected for inhalation or for delivery through the digestive tract, such as orally. The preparation of such pharmaceutically acceptable compositions is within the skill of the art.

The formulation components are present in concentrations that are acceptable to the site of administration. For example, buffers are used to maintain the composition at physiological pH or at a slightly lower pH, typically within a pH range of from about 5 to about 8.

When parenteral administration is contemplated, the therapeutic compositions for use in this invention may be in the form of a pyrogen-free, parenterally acceptable, aqueous solution comprising the desired CHL2 molecule in a pharmaceutically acceptable vehicle. A particularly suitable vehicle for parenteral injection is sterile distilled water in which a CHL2 molecule is formulated as a sterile, isotonic solution, properly preserved. Yet another preparation can involve the formulation of the desired molecule with an agent, such as injectable microspheres, bio-erodible particles, polymeric compounds (such as polylactic acid or polyglycolic acid), beads, or liposomes, that provides for the controlled or sustained release of the product which may then be delivered via a depot injection. Hyaluronic acid may also be used, and this may have the effect of promoting sustained duration in the circulation. Other suitable means for the introduction of the desired molecule include implantable drug delivery devices.

In one embodiment, a pharmaceutical composition may be formulated for inhalation. For example, CHL2 polypeptide may be formulated as a dry powder for inhalation. CHL2 polypeptide or nucleic acid molecule inhalation solutions may also be formulated with a propellant for aerosol delivery. In yet another embodiment, solutions may be nebulized. Pulmonary administration is further described in PCT Pub. No. WO 94/20069, which describes the pulmonary delivery of chemically modified proteins.

It is also contemplated that certain formulations may be administered orally. In one embodiment of the present invention, CHL2 polypeptides that are administered in this fashion can be formulated with or without those carriers customarily used in the compounding of solid dosage forms such as tablets and capsules. For example, a capsule may be designed to release the active portion of the formulation at the point in the gastrointestinal tract when bioavailability is maximized and pre-systemic degradation is minimized. Additional agents can be included to facilitate absorption of the CHL2 polypeptide. Diluents, flavorings, low melting point waxes, vegetable oils, lubricants, suspending agents, tablet disintegrating agents, and binders may also be employed.

Another pharmaceutical composition may involve an effective quantity of CHL2 polypeptides in a mixture with non-toxic excipients that are suitable for the manufacture of tablets. By dissolving the tablets in sterile water, or another appropriate vehicle, solutions can be prepared in unit-dose form. Suitable excipients include, but are not limited to, inert diluents, such as calcium carbonate, sodium carbonate or bicarbonate, lactose, or calcium phosphate; or binding agents, such as starch, gelatin, or acacia; or lubricating agents such as magnesium stearate, stearic acid, or talc.

Additional CHL2 polypeptide pharmaceutical compositions will be evident to those skilled in the art, including formulations involving CHL2 polypeptides in sustained- or controlled-delivery formulations. Techniques for formulating a variety of other sustained- or controlled-delivery means, such as liposome carriers, bio-erodible microparticles or porous beads and depot injections, are also known to those skilled in the art. *See, e.g.*, PCT/US93/00829, which describes the controlled release of porous polymeric microparticles for the delivery of pharmaceutical compositions.

Additional examples of sustained-release preparations include semipermeable polymer matrices in the form of shaped articles, *e.g.* films, or microcapsules. Sustained release matrices may include polyesters, hydrogels, polylactides (U.S. Patent No. 3,773,919 and European Patent No. 058481), copolymers of L-glutamic acid and gamma ethyl-L-glutamate (Sidman *et al.*,

1983, *Biopolymers* 22:547-56), poly(2-hydroxyethyl-methacrylate) (Langer *et al.*, 1981, *J. Biomed. Mater. Res.* 15:167-277 and Langer, 1982, *Chem. Tech.* 12:98-105), ethylene vinyl acetate (Langer *et al.*, *supra*) or poly-D(-)-3-hydroxybutyric acid (European Patent No. 133988). Sustained-release compositions may also include liposomes, which can be prepared by any of several methods known in the art. *See, e.g.*, Eppstein *et al.*, 1985, *Proc. Natl. Acad. Sci. USA* 82:3688-92; and European Patent Nos. 036676, 088046, and 143949.

The CHL2 pharmaceutical composition to be used for *in vivo* administration typically must be sterile. This may be accomplished by filtration through sterile filtration membranes. Where the composition is lyophilized, sterilization using this method may be conducted either prior to, or following, lyophilization and reconstitution. The composition for parenteral administration may be stored in lyophilized form or in a solution. In addition, parenteral compositions generally are placed into a container having a sterile access port, for example, an intravenous solution bag or vial having a stopper pierceable by a hypodermic injection needle.

Once the pharmaceutical composition has been formulated, it may be stored in sterile vials as a solution, suspension, gel, emulsion, solid, or as a dehydrated or lyophilized powder. Such formulations may be stored either in a ready-to-use form or in a form (*e.g.*, lyophilized) requiring reconstitution prior to administration.

In a specific embodiment, the present invention is directed to kits for producing a single-dose administration unit. The kits may each contain both a first container having a dried protein and a second container having an aqueous formulation. Also included within the scope of this invention are kits containing single and multi-chambered pre-filled syringes (*e.g.*, liquid syringes and lyosyringes).

The effective amount of a CHL2 pharmaceutical composition to be employed therapeutically will depend, for example, upon the therapeutic context and objectives. One skilled in the art will appreciate that the appropriate dosage levels for treatment will thus vary depending, in part, upon the molecule

delivered, the indication for which the CHL2 molecule is being used, the route of administration, and the size (body weight, body surface, or organ size) and condition (the age and general health) of the patient. Accordingly, the clinician may titer the dosage and modify the route of administration to obtain the optimal therapeutic effect. A typical dosage may range from about 0.1 μg/kg to up to about 100 mg/kg or more, depending on the factors mentioned above. In other embodiments, the dosage may range from 0.1 μg/kg up to about 100 mg/kg; or 1 μg/kg up to about 100 mg/kg; or 5 μg/kg up to about 100 mg/kg.

The frequency of dosing will depend upon the pharmacokinetic parameters of the CHL2 molecule in the formulation being used. Typically, a clinician will administer the composition until a dosage is reached that achieves the desired effect. The composition may therefore be administered as a single dose, as two or more doses (which may or may not contain the same amount of the desired molecule) over time, or as a continuous infusion via an implantation device or catheter. Further refinement of the appropriate dosage is routinely made by those of ordinary skill in the art and is within the ambit of tasks routinely performed by them. Appropriate dosages may be ascertained through use of appropriate dose-response data.

The route of administration of the pharmaceutical composition is in accord with known methods, *e.g.*, orally; through injection by intravenous, intraperitoneal, intracerebral (intraparenchymal), intracerebroventricular, intramuscular, intraocular, intraarterial, intraportal, or intralesional routes; by sustained release systems; or by implantation devices. Where desired, the compositions may be administered by bolus injection or continuously by infusion, or by implantation device.

Alternatively or additionally, the composition may be administered locally via implantation of a membrane, sponge, or other appropriate material onto which the desired molecule has been absorbed or encapsulated. Where an implantation device is used, the device may be implanted into any suitable tissue or organ, and delivery of the desired molecule may be via diffusion, timed-release bolus, or continuous administration.

In some cases, it may be desirable to use CHL2 polypeptide pharmaceutical compositions in an *ex vivo* manner. In such instances, cells, tissues, or organs that have been removed from the patient are exposed to CHL2 polypeptide pharmaceutical compositions after which the cells, tissues, or organs are subsequently implanted back into the patient.

In other cases, a CHL2 polypeptide can be delivered by implanting certain cells that have been genetically engineered, using methods such as those described herein, to express and secrete the CHL2 polypeptide. Such cells may be animal or human cells, and may be autologous, heterologous, or xenogeneic. Optionally, the cells may be immortalized. In order to decrease the chance of an immunological response, the cells may be encapsulated to avoid infiltration of surrounding tissues. The encapsulation materials are typically biocompatible, semi-permeable polymeric enclosures or membranes that allow the release of the protein product(s) but prevent the destruction of the cells by the patient's immune system or by other detrimental factors from the surrounding tissues.

As discussed herein, it may be desirable to treat isolated cell populations (such as stem cells, lymphocytes, red blood cells, chondrocytes, neurons, and the like) with one or more CHL2 polypeptides. This can be accomplished by exposing the isolated cells to the polypeptide directly, where it is in a form that is permeable to the cell membrane.

Additional embodiments of the present invention relate to cells and methods (*e.g.*, homologous recombination and/or other recombinant production methods) for both the *in vitro* production of therapeutic polypeptides and for the production and delivery of therapeutic polypeptides by gene therapy or cell therapy. Homologous and other recombination methods may be used to modify a cell that contains a normally transcriptionally-silent CHL2 gene, or an under-expressed gene, and thereby produce a cell which expresses therapeutically efficacious amounts of CHL2 polypeptides.

Homologous recombination is a technique originally developed for targeting genes to induce or correct mutations in transcriptionally active genes. Kucherlapati, 1989, *Prog. in Nucl. Acid Res. & Mol. Biol.* 36:301. The basic

technique was developed as a method for introducing specific mutations into specific regions of the mammalian genome (Thomas *et al.*, 1986, *Cell* 44:419-28; Thomas and Capecchi, 1987, *Cell* 51:503-12; Doetschman *et al.*, 1988, *Proc. Natl. Acad. Sci. U.S.A.* 85:8583-87) or to correct specific mutations within defective genes (Doetschman *et al.*, 1987, *Nature* 330:576-78). Exemplary homologous recombination techniques are described in U.S. Patent No. 5,272,071; European Patent Nos. 9193051 and 505500; PCT/US90/07642, and PCT Pub No. WO 91/09955).

Through homologous recombination, the DNA sequence to be inserted into the genome can be directed to a specific region of the gene of interest by attaching it to targeting DNA. The targeting DNA is a nucleotide sequence that is complementary (homologous) to a region of the genomic DNA. Small pieces of targeting DNA that are complementary to a specific region of the genome are put in contact with the parental strand during the DNA replication process. It is a general property of DNA that has been inserted into a cell to hybridize, and therefore, recombine with other pieces of endogenous DNA through shared homologous regions. If this complementary strand is attached to an oligonucleotide that contains a mutation or a different sequence or an additional nucleotide, it too is incorporated into the newly synthesized strand as a result of the recombination. As a result of the proofreading function, it is possible for the new sequence of DNA to serve as the template. Thus, the transferred DNA is incorporated into the genome.

Attached to these pieces of targeting DNA are regions of DNA that may interact with or control the expression of a CHL2 polypeptide, *e.g.*, flanking sequences. For example, a promoter/enhancer element, a suppressor, or an exogenous transcription modulatory element is inserted in the genome of the intended host cell in proximity and orientation sufficient to influence the transcription of DNA encoding the desired CHL2 polypeptide. The control element controls a portion of the DNA present in the host cell genome. Thus, the expression of the desired CHL2 polypeptide may be achieved not by transfection of DNA that encodes the CHL2 gene itself, but rather by the use of targeting DNA

(containing regions of homology with the endogenous gene of interest) coupled with DNA regulatory segments that provide the endogenous gene sequence with recognizable signals for transcription of a CHL2 gene.

In an exemplary method, the expression of a desired targeted gene in a cell (*i.e.*, a desired endogenous cellular gene) is altered via homologous recombination into the cellular genome at a preselected site, by the introduction of DNA which includes at least a regulatory sequence, an exon, and a splice donor site. These components are introduced into the chromosomal (genomic) DNA in such a manner that this, in effect, results in the production of a new transcription unit (in which the regulatory sequence, the exon, and the splice donor site present in the DNA construct are operatively linked to the endogenous gene). As a result of the introduction of these components into the chromosomal DNA, the expression of the desired endogenous gene is altered.

Altered gene expression, as described herein, encompasses activating (or causing to be expressed) a gene which is normally silent (unexpressed) in the cell as obtained, as well as increasing the expression of a gene which is not expressed at physiologically significant levels in the cell as obtained. The embodiments further encompass changing the pattern of regulation or induction such that it is different from the pattern of regulation or induction that occurs in the cell as obtained, and reducing (including eliminating) the expression of a gene which is expressed in the cell as obtained.

One method by which homologous recombination can be used to increase, or cause, CHL2 polypeptide production from a cell's endogenous CHL2 gene involves first using homologous recombination to place a recombination sequence from a site-specific recombination system (*e.g.*, Cre/loxP, FLP/FRT) (Sauer, 1994, *Curr. Opin. Biotechnol.*, 5:521-27; Sauer, 1993, *Methods Enzymol.*, 225:890-900) upstream of (*i.e.*, 5' to) the cell's endogenous genomic CHL2 polypeptide coding region. A plasmid containing a recombination site homologous to the site that was placed just upstream of the genomic CHL2 polypeptide coding region is introduced into the modified cell line along with the appropriate recombinase enzyme. This recombinase causes the plasmid to

integrate, via the plasmid's recombination site, into the recombination site located just upstream of the genomic CHL2 polypeptide coding region in the cell line (Baubonis and Sauer, 1993, *Nucleic Acids Res.* 21:2025-29; O'Gorman *et al.*, 1991, *Science* 251:1351-55). Any flanking sequences known to increase transcription (*e.g.*, enhancer/promoter, intron, translational enhancer), if properly positioned in this plasmid, would integrate in such a manner as to create a new or modified transcriptional unit resulting in *de novo* or increased CHL2 polypeptide production from the cell's endogenous CHL2 gene.

A further method to use the cell line in which the site specific recombination sequence had been placed just upstream of the cell's endogenous genomic CHL2 polypeptide coding region is to use homologous recombination to introduce a second recombination site elsewhere in the cell line's genome. The appropriate recombinase enzyme is then introduced into the two-recombination-site cell line, causing a recombination event (deletion, inversion, and translocation) (Sauer, 1994, *Curr. Opin. Biotechnol.*, 5:521-27; Sauer, 1993, *Methods Enzymol.*, 225:890-900) that would create a new or modified transcriptional unit resulting in *de novo* or increased CHL2 polypeptide production from the cell's endogenous CHL2 gene.

An additional approach for increasing, or causing, the expression of CHL2 polypeptide from a cell's endogenous CHL2 gene involves increasing, or causing, the expression of a gene or genes (*e.g.*, transcription factors) and/or decreasing the expression of a gene or genes (*e.g.*, transcriptional repressors) in a manner which results in *de novo* or increased CHL2 polypeptide production from the cell's endogenous CHL2 gene. This method includes the introduction of a non-naturally occurring polypeptide (*e.g.*, a polypeptide comprising a site specific DNA binding domain fused to a transcriptional factor domain) into the cell such that *de novo* or increased CHL2 polypeptide production from the cell's endogenous CHL2 gene results.

The present invention further relates to DNA constructs useful in the method of altering expression of a target gene. In certain embodiments, the exemplary DNA constructs comprise: (a) one or more targeting sequences, (b) a

regulatory sequence, (c) an exon, and (d) an unpaired splice-donor site. The targeting sequence in the DNA construct directs the integration of elements (a) - (d) into a target gene in a cell such that the elements (b) - (d) are operatively linked to sequences of the endogenous target gene. In another embodiment, the DNA constructs comprise: (a) one or more targeting sequences, (b) a regulatory sequence, (c) an exon, (d) a splice-donor site, (e) an intron, and (f) a splice-acceptor site, wherein the targeting sequence directs the integration of elements (a) - (f) such that the elements of (b) - (f) are operatively linked to the endogenous gene. The targeting sequence is homologous to the preselected site in the cellular chromosomal DNA with which homologous recombination is to occur. In the construct, the exon is generally 3' of the regulatory sequence and the splice-donor site is 3' of the exon.

If the sequence of a particular gene is known, such as the nucleic acid sequence of CHL2 polypeptide presented herein, a piece of DNA that is complementary to a selected region of the gene can be synthesized or otherwise obtained, such as by appropriate restriction of the native DNA at specific recognition sites bounding the region of interest. This piece serves as a targeting sequence upon insertion into the cell and will hybridize to its homologous region within the genome. If this hybridization occurs during DNA replication, this piece of DNA, and any additional sequence attached thereto, will act as an Okazaki fragment and will be incorporated into the newly synthesized daughter strand of DNA. The present invention, therefore, includes nucleotides encoding a CHL2 polypeptide, which nucleotides may be used as targeting sequences.

CHL2 polypeptide cell therapy, *e.g.*, the implantation of cells producing CHL2 polypeptides, is also contemplated. This embodiment involves implanting cells capable of synthesizing and secreting a biologically active form of CHL2 polypeptide. Such CHL2 polypeptide-producing cells can be cells that are natural producers of CHL2 polypeptides or may be recombinant cells whose ability to produce CHL2 polypeptides has been augmented by transformation with a gene encoding the desired CHL2 polypeptide or with a gene augmenting the expression of CHL2 polypeptide. Such a modification may be accomplished by means of a

vector suitable for delivering the gene as well as promoting its expression and secretion. In order to minimize a potential immunological reaction in patients being administered a CHL2 polypeptide, as may occur with the administration of a polypeptide of a foreign species, it is preferred that the natural cells producing CHL2 polypeptide be of human origin and produce human CHL2 polypeptide. Likewise, it is preferred that the recombinant cells producing CHL2 polypeptide be transformed with an expression vector containing a gene encoding a human CHL2 polypeptide.

Implanted cells may be encapsulated to avoid the infiltration of surrounding tissue. Human or non-human animal cells may be implanted in patients in biocompatible, semipermeable polymeric enclosures or membranes that allow the release of CHL2 polypeptide, but that prevent the destruction of the cells by the patient's immune system or by other detrimental factors from the surrounding tissue. Alternatively, the patient's own cells, transformed to produce CHL2 polypeptides *ex vivo*, may be implanted directly into the patient without such encapsulation.

Techniques for the encapsulation of living cells are known in the art, and the preparation of the encapsulated cells and their implantation in patients may be routinely accomplished. For example, Baetge *et al.* (PCT Pub. No. WO 95/05452 and PCT/US94/09299) describe membrane capsules containing genetically engineered cells for the effective delivery of biologically active molecules. The capsules are biocompatible and are easily retrievable. The capsules encapsulate cells transfected with recombinant DNA molecules comprising DNA sequences coding for biologically active molecules operatively linked to promoters that are not subject to down-regulation *in vivo* upon implantation into a mammalian host. The devices provide for the delivery of the molecules from living cells to specific sites within a recipient. In addition, *see* U.S. Patent Nos. 4,892,538; 5,011,472; and 5,106,627. A system for encapsulating living cells is described in PCT Pub. No. WO 91/10425 (Aebischer *et al.*). *See also*, PCT Pub. No. WO 91/10470 (Aebischer *et al.*); Winn *et al.*, 1991, *Exper. Neurol.* 113:322-29; Aebischer *et al.*, 1991, *Exper. Neurol.* 111:269-75; and Tresco *et al.*, 1992, *ASAIO* 38:17-23.

*In vivo* and *in vitro* gene therapy delivery of CHL2 polypeptides is also envisioned. One example of a gene therapy technique is to use the CHL2 gene (either genomic DNA, cDNA, and/or synthetic DNA) encoding a CHL2 polypeptide which may be operably linked to a constitutive or inducible promoter to form a "gene therapy DNA construct." The promoter may be homologous or heterologous to the endogenous CHL2 gene, provided that it is active in the cell or tissue type into which the construct will be inserted. Other components of the gene therapy DNA construct may optionally include DNA molecules designed for site-specific integration (*e.g.*, endogenous sequences useful for homologous recombination), tissue-specific promoters, enhancers or silencers, DNA molecules capable of providing a selective advantage over the parent cell, DNA molecules useful as labels to identify transformed cells, negative selection systems, cell specific binding agents (as, for example, for cell targeting), cell-specific internalization factors, transcription factors enhancing expression from a vector, and factors enabling vector production.

A gene therapy DNA construct can then be introduced into cells (either *ex vivo* or *in vivo*) using viral or non-viral vectors. One means for introducing the gene therapy DNA construct is by means of viral vectors as described herein. Certain vectors, such as retroviral vectors, will deliver the DNA construct to the chromosomal DNA of the cells, and the gene can integrate into the chromosomal DNA. Other vectors will function as episomes, and the gene therapy DNA construct will remain in the cytoplasm.

In yet other embodiments, regulatory elements can be included for the controlled expression of the CHL2 gene in the target cell. Such elements are turned on in response to an appropriate effector. In this way, a therapeutic polypeptide can be expressed when desired. One conventional control means involves the use of small molecule dimerizers or rapalogs to dimerize chimeric proteins which contain a small molecule-binding domain and a domain capable of initiating a biological process, such as a DNA-binding protein or transcriptional activation protein (*see* PCT Pub. Nos. WO 96/41865, WO 97/31898, and WO

97/31899). The dimerization of the proteins can be used to initiate transcription of the transgene.

An alternative regulation technology uses a method of storing proteins expressed from the gene of interest inside the cell as an aggregate or cluster. The gene of interest is expressed as a fusion protein that includes a conditional aggregation domain that results in the retention of the aggregated protein in the endoplasmic reticulum. The stored proteins are stable and inactive inside the cell. The proteins can be released, however, by administering a drug (*e.g.*, small molecule ligand) that removes the conditional aggregation domain and thereby specifically breaks apart the aggregates or clusters so that the proteins may be secreted from the cell. *See* Aridor *et al.*, 2000, *Science* 287:816-17 and Rivera *et al.*, 2000, *Science* 287:826-30.

Other suitable control means or gene switches include, but are not limited to, the systems described herein. Mifepristone (RU486) is used as a progesterone antagonist. The binding of a modified progesterone receptor ligand-binding domain to the progesterone antagonist activates transcription by forming a dimer of two transcription factors that then pass into the nucleus to bind DNA. The ligand-binding domain is modified to eliminate the ability of the receptor to bind to the natural ligand. The modified steroid hormone receptor system is further described in U.S. Patent No. 5,364,791 and PCT Pub. Nos. WO 96/40911 and WO 97/10337.

Yet another control system uses ecdysone (a fruit fly steroid hormone) which binds to and activates an ecdysone receptor (cytoplasmic receptor). The receptor then translocates to the nucleus to bind a specific DNA response element (promoter from ecdysone-responsive gene). The ecdysone receptor includes a transactivation domain, DNA-binding domain, and ligand-binding domain to initiate transcription. The ecdysone system is further described in U.S. Patent No. 5,514,578 and PCT Pub. Nos. WO 97/38117, WO 96/37609, and WO 93/03162.

Another control means uses a positive tetracycline-controllable transactivator. This system involves a mutated tet repressor protein DNA-binding domain (mutated tet R-4 amino acid changes which resulted in a reverse

tetracycline-regulated transactivator protein, *i.e.*, it binds to a tet operator in the presence of tetracycline) linked to a polypeptide which activates transcription. Such systems are described in U.S. Patent Nos. 5,464,758, 5,650,298, and 5,654,168.

Additional expression control systems and nucleic acid constructs are described in U.S. Patent Nos. 5,741,679 and 5,834,186, to Innovir Laboratories Inc.

*In vivo* gene therapy may be accomplished by introducing the gene encoding CHL2 polypeptide into cells via local injection of a CHL2 nucleic acid molecule or by other appropriate viral or non-viral delivery vectors. Hefti 1994, *Neurobiology* 25:1418-35. For example, a nucleic acid molecule encoding a CHL2 polypeptide may be contained in an adeno-associated virus (AAV) vector for delivery to the targeted cells (*see, e.g.*, Johnson, PCT Pub. No. WO 95/34670; PCT App. No. PCT/US95/07178). The recombinant AAV genome typically contains AAV inverted terminal repeats flanking a DNA sequence encoding a CHL2 polypeptide operably linked to functional promoter and polyadenylation sequences.

Alternative suitable viral vectors include, but are not limited to, retrovirus, adenovirus, herpes simplex virus, lentivirus, hepatitis virus, parvovirus, papovavirus, poxvirus, alphavirus, coronavirus, rhabdovirus, paramyxovirus, and papilloma virus vectors. U.S. Patent No. 5,672,344 describes an *in vivo* viral-mediated gene transfer system involving a recombinant neurotrophic HSV-1 vector. U.S. Patent No. 5,399,346 provides examples of a process for providing a patient with a therapeutic protein by the delivery of human cells which have been treated *in vitro* to insert a DNA segment encoding a therapeutic protein. Additional methods and materials for the practice of gene therapy techniques are described in U.S. Patent Nos. 5,631,236 (involving adenoviral vectors), 5,672,510 (involving retroviral vectors), 5,635,399 (involving retroviral vectors expressing cytokines).

Nonviral delivery methods include, but are not limited to, liposome-mediated transfer, naked DNA delivery (direct injection), receptor-mediated

transfer (ligand-DNA complex), electroporation, calcium phosphate precipitation, and microparticle bombardment (*e.g.*, gene gun). Gene therapy materials and methods may also include inducible promoters, tissue-specific enhancer-promoters, DNA sequences designed for site-specific integration, DNA sequences capable of providing a selective advantage over the parent cell, labels to identify transformed cells, negative selection systems and expression control systems (safety measures), cell-specific binding agents (for cell targeting), cell-specific internalization factors, and transcription factors to enhance expression by a vector as well as methods of vector manufacture. Such additional methods and materials for the practice of gene therapy techniques are described in U.S. Patent Nos. 4,970,154 (involving electroporation techniques), 5,679,559 (describing a lipoprotein-containing system for gene delivery), 5,676,954 (involving liposome carriers), 5,593,875 (describing methods for calcium phosphate transfection), and 4,945,050 (describing a process wherein biologically active particles are propelled at cells at a speed whereby the particles penetrate the surface of the cells and become incorporated into the interior of the cells), and PCT Pub. No. WO 96/40958 (involving nuclear ligands).

It is also contemplated that CHL2 gene therapy or cell therapy can further include the delivery of one or more additional polypeptide(s) in the same or a different cell(s). Such cells may be separately introduced into the patient, or the cells may be contained in a single implantable device, such as the encapsulating membrane described above, or the cells may be separately modified by means of viral vectors.

A means to increase endogenous CHL2 polypeptide expression in a cell via gene therapy is to insert one or more enhancer elements into the CHL2 polypeptide promoter, where the enhancer elements can serve to increase transcriptional activity of the CHL2 gene. The enhancer elements used will be selected based on the tissue in which one desires to activate the gene – enhancer elements known to confer promoter activation in that tissue will be selected. For example, if a gene encoding a CHL2 polypeptide is to be "turned on" in T-cells, the *lck* promoter enhancer element may be used. Here, the functional portion of

the transcriptional element to be added may be inserted into a fragment of DNA containing the CHL2 polypeptide promoter (and optionally, inserted into a vector and/or 5' and/or 3' flanking sequences) using standard cloning techniques. This construct, known as a "homologous recombination construct," can then be introduced into the desired cells either *ex vivo* or *in vivo*.

Gene therapy also can be used to decrease CHL2 polypeptide expression by modifying the nucleotide sequence of the endogenous promoter. Such modification is typically accomplished via homologous recombination methods. For example, a DNA molecule containing all or a portion of the promoter of the CHL2 gene selected for inactivation can be engineered to remove and/or replace pieces of the promoter that regulate transcription. For example, the TATA box and/or the binding site of a transcriptional activator of the promoter may be deleted using standard molecular biology techniques; such deletion can inhibit promoter activity thereby repressing the transcription of the corresponding CHL2 gene. The deletion of the TATA box or the transcription activator binding site in the promoter may be accomplished by generating a DNA construct comprising all or the relevant portion of the CHL2 polypeptide promoter (from the same or a related species as the CHL2 gene to be regulated) in which one or more of the TATA box and/or transcriptional activator binding site nucleotides are mutated via substitution, deletion and/or insertion of one or more nucleotides. As a result, the TATA box and/or activator binding site has decreased activity or is rendered completely inactive. This construct, which also will typically contain at least about 500 bases of DNA that correspond to the native (endogenous) 5' and 3' DNA sequences adjacent to the promoter segment that has been modified, may be introduced into the appropriate cells (either *ex vivo* or *in vivo*) either directly or via a viral vector as described herein. Typically, the integration of the construct into the genomic DNA of the cells will be via homologous recombination, where the 5' and 3' DNA sequences in the promoter construct can serve to help integrate the modified promoter region via hybridization to the endogenous chromosomal DNA.

Therapeutic Uses

CHL2 nucleic acid molecules, polypeptides, and agonists and antagonists thereof can be used to treat, diagnose, ameliorate, or prevent a number of diseases, disorders, or conditions, including those recited herein.

CHL2 polypeptide agonists and antagonists include those molecules which regulate CHL2 polypeptide activity and either increase or decrease at least one activity of the mature form of the CHL2 polypeptide. Agonists or antagonists may be co-factors, such as a protein, peptide, carbohydrate, lipid, or small molecular weight molecule, which interact with CHL2 polypeptide and thereby regulate its activity. Potential polypeptide agonists or antagonists include antibodies that react with either soluble or membrane-bound forms of CHL2 polypeptides that comprise part or all of the extracellular domains of the said proteins. Molecules that regulate CHL2 polypeptide expression typically include nucleic acids encoding CHL2 polypeptide that can act as anti-sense regulators of expression.

One of the major roles of the BMP-family of gene products, specifically BMP2 and BMP4, is the regulation of bone-mass in the adult. Since BMP1 has been shown to cleave and inactivate CHD, and has been isolated with BMP2 and BMP3 from bone (Wozney *et al.*, 1988, *Science* 242:1528-34; Celeste *et al.*, 1990, *Proc. Nat. Acad. Sci. USA* 87: 9843-47), CHL2 polypeptides may play a key regulatory role in osteogenesis. Accordingly, CHL2 nucleic acid molecules, polypeptides, and agonists and antagonists thereof (including, but not limited to, anti-CHL2 selective binding agents) may be useful in diagnosing or treating diseases and conditions affecting bone density. Examples of such diseases and conditions include, but are not limited to, osteopetrosis and osteoporosis. Other diseases and conditions affecting bone density are encompassed within the scope of this invention.

The direct delivery of BMP4 or other BMP-family members to the regenerating bone through the blood stream appears to be a straightforward therapeutic concept for treatment of osteopetrosis. However, it may be difficult to accomplish since BMP4 is known to travel only a short distance *in vivo* (Jones *et*

*al.*, 1996, *Curr. Biol.* 6:1468-75). As is the case with CHD during embryogenesis, the formation of a complex between BMP and CHLII polypeptide may result in the further migration of BMP and formation of a BMP concentration gradient (Jones and Smith, 1998, *Dev. Biol.* 194:12-17).

Based on the tight spatial regulation of CHL2 gene expression at the surface of the articular cartilage – where the first sign of cartilage damage is detected during the pathogenesis of osteoarthritis – changes of expression levels of CHL2 polypeptide may play a role in the pathogenesis of osteoarthritis or rheumatoid arthritis. Accordingly, CHL2 nucleic acid molecules, polypeptides, and agonists and antagonists thereof may be useful in preventing cartilage fibrosis during the early phase of osteoarthritis, regenerating the once-disrupted superficial zone in the later phases of osteoarthritis or rheumatoid arthritis, or generating a proper superficial zone (i.e. surface) in transplanted cartilages.

BMP polypeptides have been shown to function in organ formation during late embryogenesis. It has been shown that organ formation in embryonic kidney, lung, and gut are affected by BMP4 expression (Hogan, 1996, *Genes Dev.* 10:1580-94). A combination of BMP4 and CHL2 polypeptides may be useful for controlling the proliferation and differentiation of progenitor cells, thus permitting the regulation of tissue regeneration or wound healing *in vivo*. Accordingly, CHL2 nucleic acid molecules, polypeptides, and agonists and antagonists thereof may be useful in promoting tissue regeneration or wound healing.

The CHL2 nucleic acid molecules, polypeptides, agonists and antagonists thereof may also be used in hematopoietic stem cell-genesis and expansion. It may be possible to use BMP/CHL2 polypeptide complexes to regulate primitive hematopoietic stem cells and thereby control adult-marrow repopulating stem cells. Alternatively, it may be possible to control stem cell genesis from a mesodermal stem cell. Thus, the CHL2 polypeptides and nucleic acids of the present invention, along with BMP, may be useful for *ex vivo* expansion of hematopoietic stem cells and gene therapy performed through such cells.

BMP4 is an essential factor for generating hematopoietic progenitor cells from the mouse ES cells. However, the effective concentration of BMP4 falls

into a narrow range (0.5 ng/ml to 5 ng/ml), which is consistent with the idea that the difference in the active concentration of BMP correlates with the difference in the resulting cell-type from the totipotent epiblast. A system for the reproducible *in vitro* generation of hematopoietic stem cells from ES cells has not yet been disclosed. However, it may be achieved by precise control of the concentration of BMP4. The CHL2 nucleic acid molecules and polypeptides of the present invention may be useful in optimizing the culturing conditions for the *in vitro* generation of hematopoietic stem cells from ES cells.

Primitive hematopoietic stem cells have been recently defined in the mouse yolk sac (Yoder *et al.*, 1997, *Proc. Natl. Acad. Sci. USA* 94:6776-80). This subclass of hematopoietic stem cells does not exhibit a marrow-repopulating activity in adults. However, when exposed to a newborn liver environment, the primitive stem cells are converted to long-term marrow-repopulating stem cells (*i.e.*, definitive stem cells). This can also be accomplished by culturing the primitive stem cells on certain stroma cell lines. The long-term survival of the definitive stem cells in culture and the long-term maintenance of the primitive stem cells that are spontaneously differentiated into definitive stem cells have not yet been established. Interactions between BMP and CHL2 polypeptide might function in the regulation of definitive stem cells. The recombinant CHL2 polypeptides and CHL2 antibodies of the present invention may be useful tools for *in vitro* long-term maintenance of hematopoietic stem cells and *in vitro* generation of definitive stem cells from primitive stem cells. Alternatively, BMP4 or putative, novel CHL2-interacting molecules may be useful for controlling these processes.

In addition, primitive hematopoietic stem cells have yet to be fully characterized. While primitive stem cells may be of a lymphocytic cell type, such cells may also be mesodermal precursors that are able to generate hematopoietic cell types as well as other mesodermal progeny. In support of this idea, adult bone marrow has recently been shown to contain endothelial progenitor cells, cells that regenerate liver (Petersen *et al.*, 1999, *Science* 284:1168-70), and a common stem cell that has a capability of deriving endothelial cells, muscle cells

and hematopoietic cells *in vivo* (Ferrari *et al.*, 1998, *Science* 279:1528-30). Furthermore, the osteoblast cell lineage, which consists of the bone marrow stroma, is known to be derived from a mesenchymal stem cell that is present in bone marrow. The possibility that a common mesodermal stem cell is responsible for the generation of both stroma and hematopoietic cells has also been previously speculated.

Since CHL2 polypeptide expression has been detected in skeletal muscle, CHL2 polypeptides may also play a role in the development and function of skeletal muscle. Accordingly, CHL2 nucleic acid molecules, polypeptides, and agonists and antagonists thereof may be useful in diagnosing or treating diseases and conditions affecting skeletal muscle. Examples of such diseases and conditions include, but are not limited to, cachexia and muscular dystrophy. Other diseases and conditions associated with skeletal muscle development and function are encompassed within the scope of this invention.

Since CHL2 polypeptide expression has been detected in the heart, CHL2 polypeptides may play a role in the development and function of the heart. Accordingly, CHL2 nucleic acid molecules, polypeptides, and agonists and antagonists thereof may also be useful in diagnosing or treating diseases and conditions affecting the heart. Examples of such diseases and conditions include, but are not limited to, arhythmias, angina, hypertension, myocardial infarction, and congestive heart failure. Other diseases and conditions associated with the heart are encompassed within the scope of this invention.

Since CHL2 polypeptide expression has been detected in the stomach, CHL2 polypeptides may play a role in the development and function of the stomach. Accordingly, CHL2 nucleic acid molecules, polypeptides, and agonists and antagonists thereof may also be useful in diagnosing or treating diseases and conditions involving the stomach. Examples of such diseases and conditions include, but are not limited to, stomach cancer and stomach ulcer. Other diseases and conditions associated with the stomach are encompassed within the scope of this invention.

Since CHL2 polypeptide expression has been detected in the liver, CHL2 polypeptides may play a role in the development and function of the liver. Accordingly, CHL2 nucleic acid molecules, polypeptides, and agonists and antagonists thereof may also be useful in diagnosing or treating diseases and conditions involving the liver. Examples of such diseases and conditions include, but are not limited to, hepatitis and hepatoma. Other diseases and conditions associated with the liver are encompassed within the scope of this invention.

Agonists or antagonists of CHL2 polypeptide function may be used (simultaneously or sequentially) in combination with one or more cytokines, growth factors, antibiotics, anti-inflammatories, and/or chemotherapeutic agents as is appropriate for the condition being treated.

Other diseases caused by or mediated by undesirable levels of CHL2 polypeptides are encompassed within the scope of the invention. Undesirable levels include excessive levels of CHL2 polypeptides and sub-normal levels of CHL2 polypeptides.

Uses of CHL2 Nucleic Acids and Polypeptides

Nucleic acid molecules of the invention (including those that do not themselves encode biologically active polypeptides) may be used to map the locations of the CHL2 gene and related genes on chromosomes. Mapping may be done by techniques known in the art, such as PCR amplification and *in situ* hybridization.

CHL2 nucleic acid molecules (including those that do not themselves encode biologically active polypeptides), may be useful as hybridization probes in diagnostic assays to test, either qualitatively or quantitatively, for the presence of a CHL2 nucleic acid molecule in mammalian tissue or bodily fluid samples.

Other methods may also be employed where it is desirable to inhibit the activity of one or more CHL2 polypeptides. Such inhibition may be effected by nucleic acid molecules that are complementary to and hybridize to expression control sequences (triple helix formation) or to CHL2 mRNA. For example, antisense DNA or RNA molecules, which have a sequence that is complementary

to at least a portion of a CHL2 gene can be introduced into the cell. Anti-sense probes may be designed by available techniques using the sequence of the CHL2 gene disclosed herein. Typically, each such antisense molecule will be complementary to the start site (5' end) of each selected CHL2 gene. When the antisense molecule then hybridizes to the corresponding CHL2 mRNA, translation of this mRNA is prevented or reduced. Anti-sense inhibitors provide information relating to the decrease or absence of a CHL2 polypeptide in a cell or organism.

Alternatively, gene therapy may be employed to create a dominant-negative inhibitor of one or more CHL2 polypeptides. In this situation, the DNA encoding a mutant polypeptide of each selected CHL2 polypeptide can be prepared and introduced into the cells of a patient using either viral or non-viral methods as described herein. Each such mutant is typically designed to compete with endogenous polypeptide in its biological role.

In addition, a CHL2 polypeptide, whether biologically active or not, may be used as an immunogen, that is, the polypeptide contains at least one epitope to which antibodies may be raised. Selective binding agents that bind to a CHL2 polypeptide (as described herein) may be used for *in vivo* and *in vitro* diagnostic purposes, including, but not limited to, use in labeled form to detect the presence of CHL2 polypeptide in a body fluid or cell sample. The antibodies may also be used to prevent, treat, or diagnose a number of diseases and disorders, including those recited herein. The antibodies may bind to a CHL2 polypeptide so as to diminish or block at least one activity characteristic of a CHL2 polypeptide, or may bind to a polypeptide to increase at least one activity characteristic of a CHL2 polypeptide (including by increasing the pharmacokinetics of the CHL2 polypeptide).

The CHL2 polypeptides of the present invention can be used to clone CHL2 polypeptide receptors, using an expression cloning strategy. Radiolabeled ($^{125}$Iodine) CHL2 polypeptide or affinity/activity-tagged CHL2 polypeptide (such as an Fc fusion or an alkaline phosphatase fusion) can be used in binding assays to identify a cell type or cell line or tissue that expresses CHL2 polypeptide receptors. RNA isolated from such cells or tissues can be converted to cDNA,

cloned into a mammalian expression vector, and transfected into mammalian cells (such as COS or 293 cells) to create an expression library. A radiolabeled or tagged CHL2 polypeptide can then be used as an affinity ligand to identify and isolate from this library the subset of cells that express the CHL2 polypeptide receptors on their surface. DNA can then be isolated from these cells and transfected into mammalian cells to create a secondary expression library in which the fraction of cells expressing CHL2 polypeptide receptors is many-fold higher than in the original library. This enrichment process can be repeated iteratively until a single recombinant clone containing a CHL2 polypeptide receptor is isolated. Isolation of the CHL2 polypeptide receptors is useful for identifying or developing novel agonists and antagonists of the CHL2 polypeptide signaling pathway. Such agonists and antagonists include soluble CHL2 polypeptide receptors, anti-CHL2 polypeptide receptor antibodies, small molecules, or antisense oligonucleotides, and they may be used for treating, preventing, or diagnosing one or more of the diseases or disorders described herein.

The murine and human CHL2 nucleic acids of the present invention are also useful tools for isolating the corresponding chromosomal CHL2 polypeptide genes. For example, mouse chromosomal DNA containing CHL2 sequences can be used to construct knockout mice, thereby permitting an examination of the *in vivo* role for CHL2 polypeptide. The human CHL2 genomic DNA can be used to identify heritable tissue-degenerating diseases.

Deposits of cDNA encoding murine and human CHL2 polypeptide, subcloned into pSPORT1 (Gibco BRL), having Accession Nos. PTA-1479 and PTA-1480, were made with the American Type Culture Collection, 10801 University Boulevard, Manassas, VA 20110-2209 on March 14, 2000.

The following examples are intended for illustration purposes only, and should not be construed as limiting the scope of the invention in any way.

Example 1: Cloning of the Murine CHL2 Polypeptide Gene

Generally, materials and methods as described in Sambrook *et al. supra* were used to clone and analyze the gene encoding murine CHL2 polypeptide.

Two murine placenta cDNA libraries were prepared in order to isolate sequences encoding murine CHL2 polypeptide. Total RNA was extracted from mouse placenta and poly-A+ RNA selected using standard extraction and isolation techniques. Random-primed cDNA was then synthesized from poly-A+ RNA using the Superscript Plasmid System for cDNA Synthesis (Gibco-BRL, Rockville, MD). The resulting cDNA was digested with Not I and fractionated on a 0.8% agarose gel. Electrophoresed cDNA of 300-1000 bp was isolated and ligated into the signal trap vector kFGF7 and cDNA of greater than 1.5 kb was isolated and ligated into pSPORT 1. Ligation reactions were introduced into *E. coli* using standard transformation techniques, transformants selected on ampicillin-containing media, and the transformants collected to generate the two cDNA libraries.

Clones containing signal peptide sequences were enriched from the kFGF-based library using kFGF signal trapping technology (U.S. Patent App. No. 09/026,959). Plasmid DNA from this cDNA library was prepared from 10 pools of 100,000 colonies each using standard techniques. This DNA was introduced into NIH 3T3 cells by calcium phosphate transfection and transfected cells were then grown for 14 days in selective media supplemented with 0.5% fetal bovine serum. Only transformants containing plasmids with signal peptide sequences generated colonies. These colonies were harvested by trypsinization and total RNA from the colonies was isolated using TRIzon reagent (Gibco-BRL) according to the manufacturer's recommended protocol. Poly-A+ RNA was isolated from the total RNA using an mRNA Purification Kit (Amersham Pharmacia Biotech, Piscataway, NJ).

To generate a cDNA library enriched for molecules containing signal peptide sequences, first strand cDNA was initially prepared from the poly-A+ RNA isolated above using the SuperScript™ preamplification system (Gibco-BRL). The first strand cDNA reaction was performed by combining 1 μg of poly-A+ RNA and 2 *p*mole of the primer 1605-21 (5'-A-A-T-C-C-G-A-T-G-C-C-C-A-

C-G-T-T-G-C-A-G-T-A-3'; SEQ ID NO: 12) in a total volume of 15 μL and heating the template-primer mixture at 70°C for 10 minutes. The template-primer mixture was transferred to 50°C, and a reaction premixture consisting of 2.5 μL of 10X buffer, 2.5 μL of 25 mM $MgCl_2$, 1.3 μL 10 mM dNTPs, and 2.5 μL 0.1 M dithiothreitol was added. Reverse transcriptase (250 U) was then added and the reaction incubated at 50°C for 1 hour. The first strand cDNA reaction was stopped by heating at 70°C for 15 minutes, and the RNA digested by treatment with 2 U of RNAse H for 20 minutes at 37°C.

Following first strand cDNA synthesis, double strand cDNA was generated by PCR amplification in a reaction containing 2 μL of first strand cDNA, the primers 1239-08 (5'-A-A-A-A-T-C-T-T-A-G-A-C-C-G-A-C-G-A-C-T-G-T-G-T-T-T-3'; SEQ ID NO: 13) and 1605-22 (5'-C-G-T-A-A-A-A-G-A-T-C-C-T-G-C-G-C-T-A-G-A-T-G-C-G-3'; SEQ ID NO: 14) at a final concentration of 0.5 μM each, 200 μM of dNTPs, and 2.5 U of Pfu polymerase (PE Biosystems, Norwalk, CT), in a volume of 100 μL. The PCR reaction was performed at 95°C for 1 minute for one cycle; 95°C for 30 seconds, 66°C for 45 seconds, and 72°C for 2 minutes for 30 cycles; and 72°C for 10 minutes for one cycle. Amplification products were digested with Not I and Sal I and ligated into kFGF7. The ligation reaction was introduced into *E. coli* using standard transformation techniques to generate a signal enriched cDNA library. Plasmid DNA was prepared from 400 selected clones and analyzed by sequencing.

The sequence of one clone (designated ymkz5-00011-c10) was found to share significant homology with the chordin precursor from *Xenopus* (GenBank accession no. Q91713). The identified clone was found to contain an insert of 418 bp encoding the N-terminal 115 amino acids of murine CHL2. A full-length cDNA clone was isolated from the cDNA library cloned into pSPORT 1 using the ymkz5-00011-c10 clone as a probe.

To screen the pSPORT 1 library, 1 x $10^6$ clones were plated on 150 mm plates at approximately 5 x $10^4$ clones per plate and the clones were then lifted from the plates on nitrocellulose filters. Filters were prehybridized in ExpressHyb hybridization solution (Clontech) for 30 minutes at 68°C and then hybridized

overnight at 68°C in the same hybridization solution containing the $^{32}$P-dCTP-labeled probe. Hybridized filters were washed twice in 2X SSC and 0.1% SDS for 10 minutes at room temperature and twice in 0.1X SSC and 0.1% SDS for 30 minutes at 65°C. Following washing, filters were subjected to autoradiography overnight at –80°C in the presence of intensifying screens. Positive clones were identified and re-screened. Two positive clones were identified following a secondary screen and DNA from these two clones was isolated and sequenced. One clone (designated pSPORTmCHL2) contained an insert of approximately 1.8 kb, which encoded the complete murine CHL2 polypeptide.

Sequence analysis of the full-length cDNA for murine CHL2 polypeptide indicated that the gene comprises a 1278 bp open reading frame encoding a protein of 426 amino acids. Figures 1A-1C illustrate the nucleotide sequence of the murine CHL2 gene (SEQ ID NO: 1) and the deduced amino acid sequence of murine CHL2 polypeptide (SEQ ID NO: 2). In Figure 1A, the signal peptide sequence, as predicted by the von Heijne algorithm, is underlined.

Computer analysis of the murine CHL2 amino acid sequence indicated that this polypeptide, like CHL (CHL1) polypeptide and CHLd5 polypeptide (co-pending and co-owned U.S. Patent App. No. 09/724,915), possesses three pro-collagen repeats (CR motifs) – in contrast to the four repeats observed in CHD (Figure 5). The amino acid sequence of CHL2 was found to share 24.4% identity (or 28% identity by a GAP search) with murine chordin. Figure 2 illustrates the amino acid sequence alignment of murine CHL2 polypeptide (mouse CHL2; SEQ ID NO: 2) and murine chordin (Af069501; SEQ ID NO: 7). The amino acid sequence of murine CHL2 polypeptide was also found to share 50% identity with CHL1 and structural similarity with BMP and SOG, specifically in the CR domains where the similarity is high. A BLAST search of the Celera human genome database indicated that the CHL2 gene shared the greatest homology with the CHL1 gene. These results suggest that the CHL2 gene is a novel member of the CHD/SOG gene family.

Example 2: Cloning of the Human CHL2 Polypeptide Gene

Generally, materials and methods as described in Sambrook *et al. supra* were used to clone and analyze the gene encoding human CHL2 polypeptide.

A human placenta cDNA library was prepared in order to isolate sequences encoding human CHL2 polypeptide. Total RNA was extracted from human placenta and poly-A+ RNA selected using standard extraction and isolation techniques. Oligo d(T)-primed cDNA was then synthesized from poly-A+ RNA using the Superscript Plasmid System for cDNA Synthesis (Gibco-BRL). The resulting cDNA was digested with Not I and fractionated on a 0.8% agarose gel. Electrophoresed cDNA of greater than 1.5 kb was isolated and ligated into pSPORT 1. Ligation reactions were introduced into *E. coli* using standard transformation techniques, transformants selected on ampicillin-containing media, and the transformants collected to generate the human placenta cDNA library.

A $^{32}$P-dCTP-labeled murine CHL2 cDNA fragment was used to screen the human placenta cDNA library. Plasmid DNA was isolated from 12 pools of 100,000 clones each, 1 μg of plasmid DNA from each pool was digested with Not I and Sal I, electrophoresed on a 0.8% agarose gel, and then transferred to a nitrocellulose filter. Filters were prehybridized in ExpressHyb hybridization solution (Clontech) for 30 minutes at 60°C and then hybridized overnight at 60°C in the same solution containing the murine CHL2 cDNA probe. Hybridized filters were washed twice in 2X SSC and 0.1% SDS for 10 minutes at room temperature and twice in 0.5X SSC and 0.1% SDS for 30 minutes at 60°C. Following washing, filters were subjected to autoradiography overnight at –80°C in the presence of intensifying screens. Positive cDNA fragments were identified in two of the twelve plasmid pools.

An individual clone containing sequences encoding human CHL2 polypeptide was isolated from the two plasmid pools identified above by plating 3 x $10^5$ clones from each pool on 150 mm plates at approximately 5 x $10^4$ clones per plate. The clones were then lifted from the plates on nitrocellulose filters and the filters analyzed as described above. Two positive clones were identified, and

each was subjected to sequence analysis. One clone (designated pSPORThCHL2) contained an insert of approximately 1.5 kb, which encoded the complete human CHL2 polypeptide.

Sequence analysis of the full-length cDNA for human CHL2 polypeptide indicated that the gene comprises a 1287 bp open reading frame encoding a protein of 429 amino acids. Figures 3A-3C illustrate the nucleotide sequence of the human CHL2 gene (SEQ ID NO: 4) and the deduced amino acid sequence of human CHL2 polypeptide (SEQ ID NO: 5). The amino acid sequence of CHL2 was found to share 26.7% identity with human chordin. Figure 4 illustrates the amino acid sequence alignment of human CHL2 polypeptide (human CHL2; SEQ ID NO: 5) and human chordin (Af076612; SEQ ID NO: 8).

Example 3: CHL2 mRNA Expression

Multiple human or murine tissue Northern blots (Clontech) were probed using a $^{32}$P-dCTP-labeled human or murine CHL2 cDNA fragment, respectively. Northern blots were prehybridized in ExpressHyb hybridization solution (Clontech) for 30 minutes at 68°C and then were hybridized in the same solution with the addition of labeled probe overnight at 68°C. Following hybridization, the filters were washed twice in 2X SSC and 0.1% SDS for 10 minutes at room temperature and twice in 0.1X SSC and 0.1% SDS for 30 minutes at 68°C. Following washing, the blots were subjected to autoradiography for 72 hours at –80°C in the presence of intensifying screens.

Northern blot analysis of human tissue blots revealed predominant CHL2 transcripts of approximately 2 kb in the prostate, testis, uterus (very abundant expression), colon, small intestine, heart, skeletal muscle, and stomach. Weak expression was detected in the trachea, placenta, and bone marrow, and very weak expression was detected in the liver. Northern blot analysis of murine tissue blots revealed predominant CHL2 transcripts of approximately 1.8 kb in liver and kidney. Very weak expression was detected in the heart. Weak expression of an approximately 2 kb transcript was detected in the testis and skeletal muscle. Expression of CHL2 in murine stomach tissue was not analyzed. Very weak

expression was detected in mouse embryonic tissue from day 7, 11, 15, and 17 embryos.

The Northern bot analysis strengthens the relationship between CHD/SOG and CHL2 polypeptide. CHD/SOG is expressed at a relatively high level in E7 embryos and at decreased levels in E11, E15, and E17 embryos (Pappano *et al.*, 1998, *Genomics* 52:236-39). CHL2 mRNA was similarly detected in E7, E11, E15, and E17 embryos, although at lower levels than are CHD/SOG transcripts. CHD/SOG is expressed in spleen, liver, and kidney (Pappano *et al.*, 1998). CHL2 mRNA was similarly detected in liver and kidney. The similarity in expression pattern, coupled with the similarity in structure, suggest that CHL2 and CHD/SOG may have similar biological activity, but that these proteins may function at different developmental stages.

The expression of CHL2 mRNA was localized by *in situ* hybridization. *In situ* hybridization to embryonic and adult mouse tissue sections was performed as described (Wilcox, 1993, *J. Histochem. Cytochem.* 41:1725-33). A murine CHL2 probe was prepared by first removing the 1.2-kb Nco I - Sal I fragment from pSPORTmCHL2 to generate pSPmCHL2COOH, which was then linearized with Eco RI. In addition, the Hind III fragment was removed to obtain pSPmCHL2NH2. Antisense-RNAs were synthesized with SP6 RNA polymerase. The RNAs synthesized from these plasmids contained non-overlapping sequences.

*In situ* hybridization analysis demonstrated that CHL2 mRNA expression began in mice at about embryonic day 12.5 in the sternum and persisted in a restricted area of adult articular cartilage. CHL2 mRNA expression was not detected in the sections examined before E10.5. At E13.5, strong CHL2 mRNA expression was detected over the chondrocytes in the limb bones (Figures 6 and 7) and sternum, as well as in the perichondrial mesenchymal cells adjacent to the developing joints. Strong but highly restricted CHL2 mRNA expression was alos detected in the superficial zone chondrocytes within the developing articular cartilage of vertebra and the epiphyses of the long bones (Figures 6 - 8). By

E18.5, and through to adulthood, the skeletal expression of CHL2 mRNA was restricted to a single layer of superficial zone chondrocytes in articular cartilage. The epiphyseal growth plate did not show CHL2 mRNA expression at any point in mouse development.

Among non-cartilagenous tissues, significant CHL2 mRNA expression were observed in ovary, oviduct, and uterus of female mice, and in testis, epididymus and possibly other accessory glands (*e.g.*, seminal vesicle, coagulating gland, and prostate) of male mice. The strong signal observed in uterine wall was somewhat analogous to the CHL1 mRNA expression detected in uterus. A weak expression of CHL2 mRNA was also detected on the colon surface. However, the location of CHL2 expression in the colon clearly differed from that of CHL1, wherein CHL2 mRNA was detected in the fibroblast/connective tissue cells dividing the submucosa and muscularis (Nakayama *et al.*, 2001). In contrast with the CHL1 gene, CHL2 mRNA expression was not detected in stomach or small intestine. Thus, among soft tissues, the CHL2 gene was expressed specifically in the reproductive organs.

Example 4: Chromosomal Mapping of the Murine CHL2 Polypeptide Gene

Fluorescence *in situ* hybridization (FISH) analysis was used to determine the chromosomal localization of the murine CHL2 gene (Shi *et al.*, 1997, *Genomics* 45:42-47). A FISH probe was prepared from a BAC clone (F1067) isolated from the Mouse ES-129/SvJ II BAC chromosome DNA library (Genome Systems) by PCR using standard techniques and primers corresponding to the 5' untranslated region of the murine CHL2 gene (5'-T-C-C-T-C-T-C-A-T-C-C-T-C-A-C-C-T-T-A-G-3'; SEQ ID NO: 15 and 5'-G-G-A-G-A-A-A-G-T-G-A-G-A-T-A-A-G-G-A-C-A-C-3'; SEQ ID NO: 16). The murine CHL2 gene was localized to chromosome 7 using the chromosome 7 centromere specific P1 clone as a co-hybridization probe (Shi *et al.*, 1997). A total of 80 metaphase cells were analyzed, and 10 of the 77 that exhibited specific labeling were used for co-hybridization experiments. As a result, the mCHL2 gene was located at a position 66% of the distance from the heterochromatic-euchromatic boundary to the

telomere of chromosome 7, which corresponds to band 7E2-E3. Thus, the CHL2 gene differs from the CHL1 gene in that the CHL2 gene is autosomal.

Example 5: Production of CHL2 Polypeptides

A. Expression of CHL2 Polypeptides in Bacteria

PCR is used to amplify template DNA sequences encoding a CHL2 polypeptide using primers corresponding to the 5' and 3' ends of the sequence. The amplified DNA products may be modified to contain restriction enzyme sites to allow for insertion into expression vectors. PCR products are gel purified and inserted into expression vectors using standard recombinant DNA methodology. An exemplary vector, such as pAMG21 (ATCC no. 98113) containing the lux promoter and a gene encoding kanamycin resistance is digested with Bam HI and Nde I for directional cloning of inserted DNA. The ligated mixture is transformed into an *E. coli* host strain by electroporation and transformants are selected for kanamycin resistance. Plasmid DNA from selected colonies is isolated and subjected to DNA sequencing to confirm the presence of the insert.

Transformed host cells are incubated in 2xYT medium containing 30 μg/mL kanamycin at 30°C prior to induction. Gene expression is induced by the addition of N-(3-oxohexanoyl)-dl-homoserine lactone to a final concentration of 30 ng/mL followed by incubation at either 30°C or 37°C for six hours. The expression of CHL2 polypeptide is evaluated by centrifugation of the culture, resuspension and lysis of the bacterial pellets, and analysis of host cell proteins by SDS-polyacrylamide gel electrophoresis.

Inclusion bodies containing CHL2 polypeptide are purified as follows. Bacterial cells are pelleted by centrifugation and resuspended in water. The cell suspension is lysed by sonication and pelleted by centrifugation at 195,000 xg for 5 to 10 minutes. The supernatant is discarded, and the pellet is washed and transferred to a homogenizer. The pellet is homogenized in 5 mL of a Percoll solution (75% liquid Percoll and 0.15 M NaCl) until uniformly suspended and then diluted and centrifuged at 21,600 xg for 30 minutes. Gradient fractions

containing the inclusion bodies are recovered and pooled. The isolated inclusion bodies are analyzed by SDS-PAGE.

A single band on an SDS polyacrylamide gel corresponding to *E. coli*-produced CHL2 polypeptide is excised from the gel, and the N-terminal amino acid sequence is determined essentially as described by Matsudaira *et al.*, 1987, *J. Biol. Chem.* 262:10-35.

B. Construction of CHL2 Polypeptide Mammalian Expression Vectors

Murine CHL2 was transiently expressed using the pSRαmCHL2 vector, which was prepared as follows. The open reading frame for CHL2 polypeptide was first amplified by PCR using standard techniques and the primers 2360-40 (5'-G-C-T-A-T-C-T-A-G-A-G-C-C-A-C-C-A-T-G-G-T-T-C-C-C-G-G-G-G-T-G-A-G-G-A-T-C-A-T-C-3'; SEQ ID NO: 17) and 2360-41 (5'-G-C-T-A-G-T-C-G-A-C-C-T-A-T-A-A-T-G-T-C-T-T-G-G-T-C-A-C-T-T-T-G-T-C-T-G-3'; SEQ ID NO: 18). The amplification product was digested with Xba I and Sal I and then inserted into an SRα-based expression plasmid (Takebe *et al.*, 1988, *Mol. Cell. Biol.* 8:466-72) to yield pSRαmCHL2.

A FLAG-tagged murine CHL2 polypeptide expression construct was prepared as follows. A full-length murine CHL2 DNA fragment, in which the stop codon was replaced by a Sal I site, was obtained by PCR using the full-length murine cDNA clone as a template and the primers 5'-G-C-T-A-G-C-G-G-C-C-G-C-G-C-C-A-C-C-A-T-G-G-T-T-C-C-C-G-G-G-G-T-G-A-G-G-A-T-C-A-T-C-3' (SEQ ID NO: 19) and 5'-G-C-T-A-G-T-C-G-A-C-T-A-A-T-G-T-C-T-T-G-G-T-C-A-C-T-T-T-G-T-C-T-G-G-G-C-3' (SEQ ID NO: 20). The amplified PCR product was digested with Not I and Sal I and then inserted into the pFLAG-CMV-5a expression vector (Sigma) with the FLAG-sequence attached in-frame with the CHL2 sequence at its carboxyl-terminus. The resulting mCHL2-FLAG expression plasmid was designated as pFLAGmCHL2.

The expression of the mCHL2-FLAG polypeptide was detected by direct western blot analysis using 10 μg/ml of the anti-FLAG mouse monoclonal antibody, M2 (Sigma), according to the manufacturer's recommendations, with

the exception that 5% (w/v) nonfat dry milk (Trader Joe's, Thousand Oaks CA) was used in place of 3% (w/v) bovine serum albumin. Rabbit polyclonal antibodies for murine CHL2 were raised using a synthetic peptide based on the carboxyl-terminal sequence: C-P-E-D-E-A-E-D-D-H-S-E-V-I-S-T-R (SEQ ID NO: 21), as described in Harlow and Lane, *Using Antibodies: A Laboratory Manual* (Cold Spring Harbor Laboratory Press, 1988). The resulting anti-sera were either used directly or subjected to affinity purification with the corresponding peptides.

To generate clones capable of stably expressing mCHL2-FLAG, 293 cells were transfected with linearized pFLAGmCHL2, stable transfectants were selected, and the expression level of the corresponding FLAG-tagged proteins in both the serum-free conditioned media and the cell lysates was compared by western blot analysis (Nakayama *et al.*, 2001, *Dev. Biol. in press*). Conditioned media was concentrated before western blotting and the proteins were then visualized. Transient transfection-based expression was carried out using 293T cells, and ten-fold concentrated conditioned media was analyzed as described above.

For mid-scale preparation, stable clones were cultured in CL 1000 flasks (Integra Biosciences, Ijamsville, MD) in 293 SFM II serum-free media and conditioned media was collected every 2-3 days. An expression level of approximately 3-4 μg/ml was obtained. The FLAG-tagged proteins were purified from 500 ml of the collected supernatant by a single-step of affinity chromatography using an anti-FLAG M2 affinity gel (Sigma) under high-salt conditions (Piccolo *et al.*, 1997, *Cell* 91:407-16). The protein concentration of each preparation was determined by western blot analysis with M2 and comparison with a FLAG-bacterial alkaline phosphatase standard (Sigma).

Example 6: Production of Anti-CHL2 Polypeptide Antibodies

Antibodies to CHL2 polypeptides may be obtained by immunization with purified protein or with CHL2 peptides produced by biological or chemical synthesis. Suitable procedures for generating antibodies include those described

in Hudson and Bay, *Practical Immunology* (2nd ed., Blackwell Scientific Publications).

In one procedure for the production of antibodies, animals (typically mice or rabbits) are injected with a CHL2 antigen (such as a CHL2 polypeptide), and those with sufficient serum titer levels as determined by ELISA are selected for hybridoma production. Spleens of immunized animals are collected and prepared as single cell suspensions from which splenocytes are recovered. The splenocytes are fused to mouse myeloma cells (such as Sp2/0-Ag14 cells), are first incubated in DMEM with 200 U/mL penicillin, 200 µg/mL streptomycin sulfate, and 4 mM glutamine, and are then incubated in HAT selection medium (hypoxanthine, aminopterin, and thymidine). After selection, the tissue culture supernatants are taken from each fusion well and tested for anti-CHL2 antibody production by ELISA.

Alternative procedures for obtaining anti-CHL2 antibodies may also be employed, such as the immunization of transgenic mice harboring human Ig loci for production of human antibodies, and the screening of synthetic antibody libraries, such as those generated by mutagenesis of an antibody variable domain.

Example 7: Biological Activity of Murine CHL2 in *Xenopus* embryos

Chordin is known to dorsalize the gastrulating *Xenopus* embryo by inhibiting the activity of BMP4. The effects of CHL2 on *Xenopus* embryo development were examined as follows. The Eco RI - Not I fragment of pSPORTmCHL2 was first cloned into pCS2+ (Rupp *et al.*, 1994, *Genes Dev.* 8:1311-23), linearized with Not I, and capped mRNAs were synthesized with SP6 polymerase and quantified (Nishinakamura *et al.*, 1999, *Dev. Biol.* 216:481-90). The pSPORTmCHL2 plasmid was also linearized directly with Not I and transcribed with T7 polymerase. Both constructs induced secondary axis formation. As a negative control, elongation factor 1 (EF1) RNA was synthesized.

Each RNA sample was injected into two ventral blastomeres of a four-cell stage *Xenopus* embryo. Following injection, the embryos were cultured in 10%

Steinberg's solution for 48 hours and were then scored for ectopic axis (Nishinakamura *et al.*, 1999; Figure 9). When blastomeres were injected with 1 *pg* of murine CHL2 RNA, the axis duplication rates ranged from 77 to 87%, whereas the rate for uninjected control embryos and EF1 RNA-injected embryos was 0%. As a positive control, experiments were also performed using murine CHL1 RNA. In these experiments, an axis duplication rate of 83% was obtained when 10-30 *pg* of CHL1 RNA was used. Thus, murine CHL2 RNA was active in antagonizing the endogenous ventralizing factor (presumably, BMP4).

Example 8: Murine CHL2 Polypeptide Inhibition of BMP4 Action

The formation of chordin-CHL1 complexes is known to inhibit BMP4 function. The similar axis duplication activity observed for CHL2 polypeptide (Example 7) suggested that CHL2 polypeptide also inhibited BMP4 action directly. $CD34^+/CD31^{hi}$ lymphohematopoietic progenitor cells, $CD34^+/CD31^{lo}$ erythro-myeloid progenitor cells, and $CD45^+$ myelomonocytic cells have been shown to be dependent on the presence of 0.15 to 2 ng/ml of BMP4 during the differentiation of mouse embryonic stem (ES) cells (Nakayama *et al.*, 2000, *Blood* 95:2275:83).

The ability of CHL2 polypeptide to inhibit BMP4 action directly was examined as follows. Rosa26 ES cells were transferred to a fibronectin-coated plate, adjusted for two days in serum-free medium, and then subjected to differentiation in serum-free methylcellulose medium in the presence of 0.9 ng/ml BMP4 (Nakayama *et al.*, 2000). A BMP4 concentration of 0.9 ng/ml resulted in a near maximal level. Embryoid bodies (EBs) were collected on day 7 and dissociated with collagenase. The cells were stained with monoclonal antibodies for hematopoietic progenitor cell markers, such as CD31 and CD34, and the stained samples were analyzed on a FACScan (Becton Dickinson, San Jose, CA).

When affinity purified mCHL2-FLAG was added over 100 ng/ml, the $CD34^-/CD31^+$ and $CD34^+/CD31^+$ cell populations were reduced to background levels (*i.e.*, levels achieved without BMP4; see Figure 10). These results suggest that the FLAG-tagged mCHL2 protein was able to inhibit the action of BMP4 *in*

*vitro.*

Example 9: Murine CHL2 Polypeptide Inhibition of BMP2 and BMP4 Action

Both BMP2 and BMP4 are known to induce alkaline phospatase expression in C2C12 myoblastic cells. CHL2 polypeptide inhibition of BMP2 and BMP4 action was demonstrated in a C2C12 alkaline phospatase assay.

The promyoblast cell line C2C12 (ATCC accession no. CRL-1772) was maintained in DMEM containing 10% fetal bovine serum (FBS) and antibiotics at 37°C in a humidified atmosphere of 5% $CO_2$. Because receptor sensitivity may decrease in cells that have been highly passaged, cells were discarded after being passaged between 10 and 15 times. C2C12 cells were plated in 96-well microtiter plates in 100 µl of DMEM containing 2% calf serum and antibiotics at a density of 3 x $10^4$ cells/well. To avoid excessive drying, the peripheral wells of the microtiter plates were filled with 200 µl of DMEM alone. Cells were then incubated overnight at 37°C in a humidified atmosphere of 5% $CO_2$.

Following plating, C2C12 cells were exposed to serial dilutions (1.47, 2.93, 5.9, 11.7, 23.4, 46.9, 93.8, 187.5, 375, 750, and 1500 ng/ml) of murine CHL2-FLAG in the presence of either 909 ng/ml (35 nM) BMP2 or 566 ng/ml (22 nM) BMP4. Cells were then incubated for three days at 37°C in a humidified atmosphere of 5% $CO_2$. Following incubation, the media was removed, the cells were rinsed with 0.1 M Tris, pH 7.4, and 150 µl of glycine buffer (0.1 M glycine, 1 mM $MgCl_2$, pH 10.5) containing 0.1% IGEPAL CA-630 (Sigma), was added to the wells. The cells were then frozen at -80°C and thawed, 50 µl of cell supernatant was removed for use in Bradford protein assays, and 100 µl of p-Nitrophenyl phosphate, disodium (Sigma; diluted to 4 mg/ml in glycine buffer) was added to the remaining cell supernatant. This mixture was incubated at 37°C for 30 minutes and 50 µl of 0.5N NaOH was added to stop the reaction. The plates were then read at 405-410 nm. CHL2-FLAG was found to inhibit both BMP2 and BMP4 action in a dose dependent manner (Figure 11).

Example 10: Direct Interaction of CHL2 Polypeptide and BMPs

The direct interaction of murine CHL1 polypeptide with human BMP4, BMP5, BMP6, and TGFβ2 has been previously demonstrated (Nakayama *et al.*, 2001). Similar immunoprecipitation experiments were performed using murine CHL2 polypeptide, with the exception that mCHL2-FLAG protein was used at 600 ng/ml. The mCHL2-FLAG protein was found to co-immunoprecipitate with BMP2, BMP4, BMP5, BMP6, GDF5 (BMP14), or activin A (Figure 12). However, a high concentration of BMP5 was required to show detectable levels of interaction with mCHL2-FLAG. CHL2 might have weaker affinities with BMP5 than with other BMPs. Similar to mCHL1, activin A – another TGFβ superfamily member – showed no sign of interaction with CHL2 under either set of conditions. However, in contrast to murine CHL1, TGFβ interaction was never observed. BMP2 and BMP4 form a separate subfamily from BMP5, BMP6, and BMP7 (Celeste *et al.*, 1990, *Proc. Natl. Acad. Sci. U.S.A.* 87:9843-47). Thus, both CHL1 and CHL2 polypeptides may be pan-BMP binding proteins.

Example 11: Expression of CHL2 Polypeptide in Transgenic Mice

To assess the biological activity of CHL2 polypeptide, a construct encoding a CHL2 polypeptide/Fc fusion protein under the control of a liver specific ApoE promoter is prepared. The delivery of this construct is expected to cause pathological changes that are informative as to the function of CHL2 polypeptide. Similarly, a construct containing the full-length CHL2 polypeptide under the control of the beta actin promoter is prepared. The delivery of this construct is expected to result in ubiquitous expression.

To generate these constructs, PCR is used to amplify template DNA sequences encoding a CHL2 polypeptide using primers that correspond to the 5' and 3' ends of the desired sequence and which incorporate restriction enzyme sites to permit insertion of the amplified product into an expression vector. Following amplification, PCR products are gel purified, digested with the appropriate restriction enzymes, and ligated into an expression vector using standard recombinant DNA techniques. For example, amplified CHL2

polypeptide sequences can be cloned into an expression vector under the control of the human β-actin promoter as described by Graham *et al.*, 1997, *Nature Genetics*, 17:272-74 and Ray *et al.*, 1991, *Genes Dev.* 5:2265-73.

Following ligation, reaction mixtures are used to transform an *E. coli* host strain by electroporation and transformants are selected for drug resistance. Plasmid DNA from selected colonies is isolated and subjected to DNA sequencing to confirm the presence of an appropriate insert and absence of mutation. The CHL2 polypeptide expression vector is purified through two rounds of CsCl density gradient centrifugation, cleaved with a suitable restriction enzyme, and the linearized fragment containing the CHL2 polypeptide transgene is purified by gel electrophoresis. The purified fragment is resuspended in 5 mM Tris, pH 7.4, and 0.2 mM EDTA at a concentration of 2 mg/mL.

Single-cell embryos from BDF1 x BDF1 bred mice are injected as described (PCT Pub. No. WO 97/23614). Embryos are cultured overnight in a $CO_2$ incubator and 15-20 two-cell embryos are transferred to the oviducts of a pseudopregnant CD1 female mice. Offspring obtained from the implantation of microinjected embryos are screened by PCR amplification of the integrated transgene in genomic DNA samples as follows. Ear pieces are digested in 20 mL ear buffer (20 mM Tris, pH 8.0, 10 mM EDTA, 0.5% SDS, and 500 mg/mL proteinase K) at 55°C overnight. The sample is then diluted with 200 mL of TE, and 2 mL of the ear sample is used in a PCR reaction using appropriate primers.

At 8 weeks of age, transgenic founder animals and control animals are sacrificed for necropsy and pathological analysis. Portions of spleen are removed and total cellular RNA isolated from the spleens using the Total RNA Extraction Kit (Qiagen) and transgene expression determined by RT-PCR. RNA recovered from spleens is converted to cDNA using the SuperScript™ Preamplification System (Gibco-BRL) as follows. A suitable primer, located in the expression vector sequence and 3' to the CHL2 polypeptide transgene, is used to prime cDNA synthesis from the transgene transcripts. Ten mg of total spleen RNA from transgenic founders and controls is incubated with 1 mM of primer for 10 minutes at 70°C and placed on ice. The reaction is then supplemented with 10 mM Tris-

HCl, pH 8.3, 50 mM KCl, 2.5 mM $MgCl_2$, 10 mM of each dNTP, 0.1 mM DTT, and 200 U of SuperScript II reverse transcriptase. Following incubation for 50 minutes at 42°C, the reaction is stopped by heating for 15 minutes at 72°C and digested with 2U of RNase H for 20 minutes at 37°C. Samples are then amplified by PCR using primers specific for CHL2 polypeptide.

Example 12: Biological Activity of CHL2 Polypeptide in Transgenic Mice

Prior to euthanasia, transgenic animals are weighed, anesthetized by isofluorane and blood drawn by cardiac puncture. The samples are subjected to hematology and serum chemistry analysis. Radiography is performed after terminal exsanguination. Upon gross dissection, major visceral organs are subject to weight analysis.

Following gross dissection, tissues (*i.e.*, liver, spleen, pancreas, stomach, the entire gastrointestinal tract, kidney, reproductive organs, skin and mammary glands, bone, brain, heart, lung, thymus, trachea, esophagus, thyroid, adrenals, urinary bladder, lymph nodes and skeletal muscle) are removed and fixed in 10% buffered Zn-Formalin for histological examination. After fixation, the tissues are processed into paraffin blocks, and 3 mm sections are obtained. All sections are stained with hematoxylin and exosin, and are then subjected to histological analysis.

The spleen, lymph node, and Peyer's patches of both the transgenic and the control mice are subjected to immunohistology analysis with B cell and T cell specific antibodies as follows. The formalin fixed paraffin embedded sections are deparaffinized and hydrated in deionized water. The sections are quenched with 3% hydrogen peroxide, blocked with Protein Block (Lipshaw, Pittsburgh, PA), and incubated in rat monoclonal anti-mouse B220 and CD3 (Harlan, Indianapolis, IN). Antibody binding is detected by biotinylated rabbit anti-rat immunoglobulins and peroxidase conjugated streptavidin (BioGenex, San Ramon, CA) with DAB as a chromagen (BioTek, Santa Barbara, CA). Sections are counterstained with hematoxylin.

After necropsy, MLN and sections of spleen and thymus from transgenic animals and control littermates are removed. Single cell suspensions are prepared by gently grinding the tissues with the flat end of a syringe against the bottom of a 100 mm nylon cell strainer (Becton Dickinson, Franklin Lakes, NJ). Cells are washed twice, counted, and approximately 1 x $10^6$ cells from each tissue are then incubated for 10 minutes with 0.5 µg CD16/32(FcγIII/II) Fc block in a 20 µL volume. Samples are then stained for 30 minutes at 2-8°C in a 100 µL volume of PBS (lacking $Ca^+$ and $Mg^+$), 0.1% bovine serum albumin, and 0.01% sodium azide with 0.5 µg antibody of FITC or PE-conjugated monoclonal antibodies against CD90.2 (Thy-1.2), CD45R (B220), CD11b(Mac-1), Gr-1, CD4, or CD8 (PharMingen, San Diego, CA). Following antibody binding, the cells are washed and then analyzed by flow cytometry on a FACScan (Becton Dickinson).

While the present invention has been described in terms of the preferred embodiments, it is understood that variations and modifications will occur to those skilled in the art. Therefore, it is intended that the appended claims cover all such equivalent variations that come within the scope of the invention as claimed.

WHAT IS CLAIMED IS:

1. An isolated nucleic acid molecule comprising a nucleotide sequence selected from the group consisting of:

(a) the nucleotide sequence as set forth in either SEQ ID NO: 1 or SEQ ID NO: 4;

(b) the nucleotide sequence of the DNA insert in either ATCC Deposit Nos. PTA-1479 or PTA-1480;

(c) a nucleotide sequence encoding the polypeptide as set forth in either SEQ ID NO: 2 or SEQ ID NO: 5;

(d) a nucleotide sequence which hybridizes under moderately or highly stringent conditions to the complement of any of (a) - (c); and

(e) a nucleotide sequence complementary to any of (a) - (c).

2. An isolated nucleic acid molecule comprising a nucleotide sequence selected from the group consisting of:

(a) a nucleotide sequence encoding a polypeptide which is at least about 70 percent identical to the polypeptide as set forth in either SEQ ID NO: 2 or SEQ ID NO: 5, wherein the encoded polypeptide has an activity of the polypeptide set forth in either SEQ ID NO: 2 or SEQ ID NO: 5;

(b) a nucleotide sequence encoding an allelic variant or splice variant of the nucleotide sequence as set forth in either SEQ ID NO: 1 or SEQ ID NO: 4, the nucleotide sequence of the DNA insert in either ATCC Deposit Nos. PTA-1479 or PTA-1480, or (a);

(c) a region of the nucleotide sequence of any of SEQ ID NO: 1 or SEQ ID NO: 4, the DNA insert in either ATCC Deposit Nos. PTA-1479 or PTA-1480, (a), or (b) encoding a polypeptide fragment of at least about 25 amino acid residues, wherein the polypeptide fragment has an activity of the encoded polypeptide as set forth in either SEQ ID NO: 2 or SEQ ID NO: 5, or is antigenic;

(d) a region of the nucleotide sequence of any of SEQ ID NO: 1 or SEQ ID NO: 4, the DNA insert in either ATCC Deposit Nos. PTA-1479 or PTA-1480, or any of (a) - (c) comprising a fragment of at least about 16 nucleotides;

(e) a nucleotide sequence which hybridizes under moderately or highly stringent conditions to the complement of any of (a) - (d); and

(f) a nucleotide sequence complementary to any of (a) - (d).

3. An isolated nucleic acid molecule comprising a nucleotide sequence selected from the group consisting of:

(a) a nucleotide sequence encoding a polypeptide as set forth in either SEQ ID NO: 2 or SEQ ID NO: 5 with at least one conservative amino acid substitution, wherein the encoded polypeptide has an activity of the polypeptide set forth in either SEQ ID NO: 2 or SEQ ID NO: 5;

(b) a nucleotide sequence encoding a polypeptide as set forth in either SEQ ID NO: 2 or SEQ ID NO: 5 with at least one amino acid insertion, wherein the encoded polypeptide has an activity of the polypeptide set forth in either SEQ ID NO: 2 or SEQ ID NO: 5;

(c) a nucleotide sequence encoding a polypeptide as set forth in either SEQ ID NO: 2 or SEQ ID NO: 5 with at least one amino acid deletion, wherein the encoded polypeptide has an activity of the polypeptide set forth in either SEQ ID NO: 2 or SEQ ID NO: 5;

(d) a nucleotide sequence encoding a polypeptide as set forth in either SEQ ID NO: 2 or SEQ ID NO: 5 which has a C- and/or N- terminal truncation, wherein the encoded polypeptide has an activity of the polypeptide set forth in either SEQ ID NO: 2 or SEQ ID NO: 5;

(e) a nucleotide sequence encoding a polypeptide as set forth in either SEQ ID NO: 2 or SEQ ID NO: 5 with at least one modification selected from the group consisting of amino acid substitutions, amino acid insertions, amino acid deletions, C-terminal truncation, and N-terminal truncation, wherein the encoded polypeptide has an activity of the polypeptide set forth in either SEQ ID NO: 2 or SEQ ID NO: 5;

(f) a nucleotide sequence of any of (a) - (e) comprising a fragment of at least about 16 nucleotides;

(g) a nucleotide sequence which hybridizes under moderately or highly stringent conditions to the complement of any of (a) - (f); and

(h) a nucleotide sequence complementary to any of (a) - (e).

4. A vector comprising the nucleic acid molecule of any of Claims 1, 2, or 3.

5. A host cell comprising the vector of Claim 4.

6. The host cell of Claim 5 that is a eukaryotic cell.

7. The host cell of Claim 5 that is a prokaryotic cell.

8. A process of producing a CHL2 polypeptide comprising culturing the host cell of Claim 5 under suitable conditions to express the polypeptide, and optionally isolating the polypeptide from the culture.

9. A polypeptide produced by the process of Claim 8.

10. The process of Claim 8, wherein the nucleic acid molecule comprises promoter DNA other than the promoter DNA for the native CHL2 polypeptide operatively linked to the DNA encoding the CHL2 polypeptide.

11. The isolated nucleic acid molecule according to Claim 2, wherein the percent identity is determined using a computer program selected from the group consisting of GAP, BLASTN, FASTA, BLASTA, BLASTX, BestFit, and the Smith-Waterman algorithm.

12. A process for determining whether a compound inhibits CHL2 polypeptide activity or CHL2 polypeptide production comprising exposing a cell according to any of Claims 5, 6, or 7 to the compound and measuring CHL2 polypeptide activity or CHL2 polypeptide production in said cell.

13. An isolated polypeptide comprising the amino acid sequence selected from the group consisting of:

(a) the amino acid sequence as set forth in either SEQ ID NO: 2 or SEQ ID NO: 5; and

(b) the amino acid sequence encoded by the DNA insert in either ATCC Deposit Nos. PTA-1479 or PTA-1480.

14. An isolated polypeptide comprising the amino acid sequence selected from the group consisting of:

(a) the amino acid sequence as set forth in either SEQ ID NO: 3 or SEQ ID NO: 6, optionally further comprising an amino-terminal methionine;

(b) an amino acid sequence for an ortholog of any of SEQ ID NO: 2 or SEQ ID NO: 5;

(c) an amino acid sequence which is at least about 70 percent identical to the amino acid sequence of any of SEQ ID NO: 2 or SEQ ID NO: 5, wherein the polypeptide has an activity of the polypeptide set forth in either SEQ ID NO: 2 or SEQ ID NO: 5;

(d) a fragment of the amino acid sequence set forth in either SEQ ID NO: 2 or SEQ ID NO: 5 comprising at least about 25 amino acid residues, wherein the fragment has an activity of the polypeptide set forth in either SEQ ID NO: 2 or SEQ ID NO: 5, or is antigenic; and

(e) an amino acid sequence for an allelic variant or splice variant of the amino acid sequence as set forth in either SEQ ID NO: 2 or SEQ ID NO: 5, the amino acid sequence encoded by the DNA insert in either ATCC Deposit Nos. PTA-1479 or PTA-1480, or any of (a) - (c).

15. An isolated polypeptide comprising the amino acid sequence selected from the group consisting of:

(a) the amino acid sequence as set forth in either SEQ ID NO: 2 or SEQ ID NO: 5 with at least one conservative amino acid substitution, wherein the polypeptide has an activity of the polypeptide set forth in either SEQ ID NO: 2 or SEQ ID NO: 5;

(b) the amino acid sequence as set forth in either SEQ ID NO: 2 or SEQ ID NO: 5 with at least one amino acid insertion, wherein the polypeptide has an activity of the polypeptide set forth in either SEQ ID NO: 2 or SEQ ID NO: 5;

(c) the amino acid sequence as set forth in either SEQ ID NO: 2 or SEQ ID NO: 5 with at least one amino acid deletion, wherein the polypeptide has an activity of the polypeptide set forth in either SEQ ID NO: 2 or SEQ ID NO: 5;

(d) the amino acid sequence as set forth in either SEQ ID NO: 2 or SEQ ID NO: 5 which has a C- and/or N- terminal truncation, wherein the polypeptide has an activity of the polypeptide set forth in either SEQ ID NO: 2 or SEQ ID NO: 5; and

(e) the amino acid sequence as set forth in either SEQ ID NO: 2 or SEQ ID NO: 5 with at least one modification selected from the group consisting of amino acid substitutions, amino acid insertions, amino acid deletions, C-terminal truncation, and N-terminal truncation, wherein the polypeptide has an activity of the polypeptide set forth in either SEQ ID NO: 2 or SEQ ID NO: 5.

16. An isolated polypeptide encoded by the nucleic acid molecule of any of Claims 1, 2, or 3, wherein the polypeptide has an activity of the polypeptide set forth in either SEQ ID NO: 2 or SEQ ID NO: 5.

17. The isolated polypeptide according to Claim 14, wherein the percent identity is determined using a computer program selected from the group consisting of GAP, BLASTP, FASTA, BLASTA, BLASTX, BestFit, and the Smith-Waterman algorithm.

18. A selective binding agent or fragment thereof which specifically binds the polypeptide of any of Claims 13, 14, or 15.

19. The selective binding agent or fragment thereof of Claim 18 that specifically binds the polypeptide comprising the amino acid sequence as set forth in either SEQ ID NO: 2 or SEQ ID NO: 5, or a fragment thereof.

20. The selective binding agent of Claim 18 that is an antibody or fragment thereof.

21. The selective binding agent of Claim 18 that is a humanized antibody.

22. The selective binding agent of Claim 18 that is a human antibody or fragment thereof.

23. The selective binding agent of Claim 18 that is a polyclonal antibody or fragment thereof.

24. The selective binding agent Claim 18 that is a monoclonal antibody or fragment thereof.

25. The selective binding agent of Claim 18 that is a chimeric antibody or fragment thereof.

26. The selective binding agent of Claim 18 that is a CDR-grafted antibody or fragment thereof.

27. The selective binding agent of Claim 18 that is an antiidiotypic antibody or fragment thereof.

28. The selective binding agent of Claim 18 that is a variable region fragment.

29. The variable region fragment of Claim 28 that is a Fab or a Fab' fragment.

30. A selective binding agent or fragment thereof comprising at least one complementarity determining region with specificity for a polypeptide having the amino acid sequence of any of SEQ ID NO: 2 or SEQ ID NO: 5.

31. The selective binding agent of Claim 18 that is bound to a detectable label.

32. The selective binding agent of Claim 18 that antagonizes CHL2 polypeptide biological activity.

33. A method for treating, preventing, or ameliorating a CHL2 polypeptide-related disease, condition, or disorder comprising administering to a patient an effective amount of a selective binding agent according to Claim 18.

34. A selective binding agent produced by immunizing an animal with a polypeptide comprising an amino acid sequence of any of SEQ ID NO: 2 or SEQ ID NO: 5.

35. A hybridoma which produces a selective binding agent which is capable of binding a polypeptide according to any of Claims 1, 2, or 3.

36. A method of detecting or quantitating the amount of CHL2 polypeptide using the selective binding agent or fragment of Claim 18.

37. A composition comprising the polypeptide of any of Claims 13, 14, or 15, and a pharmaceutically acceptable formulation agent.

38. The composition of Claim 37, wherein the pharmaceutically acceptable formulation agent is a carrier, adjuvant, solubilizer, stabilizer, or anti-oxidant.

39. The composition of Claim 37, wherein the polypeptide comprises the amino acid sequence as set forth in either SEQ ID NO: 3 or SEQ ID NO: 6.

40. A polypeptide comprising a derivative of the polypeptide of any of Claims 13, 14, or 15.

41. The polypeptide of Claim 40 that is covalently modified with a water-soluble polymer.

42. The polypeptide of Claim 41, wherein the water-soluble polymer is selected from the group consisting of polyethylene glycol, monomethoxy-polyethylene glycol, dextran, cellulose, poly-(N-vinyl pyrrolidone) polyethylene glycol, propylene glycol homopolymers, polypropylene oxide/ethylene oxide co-polymers, polyoxyethylated polyols, and polyvinyl alcohol.

43. A composition comprising a nucleic acid molecule of any of Claims 1, 2, or 3 and a pharmaceutically acceptable formulation agent.

44. The composition of Claim 43, wherein said nucleic acid molecule is contained in a viral vector.

45. A viral vector comprising a nucleic acid molecule of any of Claims 1, 2, or 3.

46. A fusion polypeptide comprising the polypeptide of any of Claims 13, 14, or 15 fused to a heterologous amino acid sequence.

47. The fusion polypeptide of Claim 46, wherein the heterologous amino acid sequence is an IgG constant domain or fragment thereof.

48. A method for treating, preventing, or ameliorating a medical condition comprising administering to a patient the polypeptide of any of Claims 13, 14, or 15, or the polypeptide encoded by the nucleic acid of any of Claims 1, 2, or 3.

49. A method of diagnosing a pathological condition or a susceptibility to a pathological condition in a subject comprising:

(a) determining the presence or amount of expression of the polypeptide of any of Claims 13, 14, or 15, or the polypeptide encoded by the nucleic acid molecule of any of Claims 1, 2, or 3 in a sample; and

(b) diagnosing a pathological condition or a susceptibility to a pathological condition based on the presence or amount of expression of the polypeptide.

50. A device, comprising:

(a) a membrane suitable for implantation; and

(b) cells encapsulated within said membrane, wherein said cells secrete a protein of any of Claims 13, 14, or 15; and

said membrane is permeable to said protein and impermeable to materials detrimental to said cells.

51. A method of identifying a compound which binds to a CHL2 polypeptide comprising:

(a) contacting the polypeptide of any of Claims 13, 14, or 15 with a compound; and

(b) determining the extent of binding of the CHL2 polypeptide to the compound.

52. The method of Claim 51, further comprising determining the activity of the polypeptide when bound to the compound.

53. A method of modulating levels of a polypeptide in an animal comprising administering to the animal the nucleic acid molecule of any of Claims 1, 2, or 3.

54. A transgenic non-human mammal comprising the nucleic acid molecule of any of Claims 1, 2, or 3.

55. A process for determining whether a compound inhibits CHL2 polypeptide activity or CHL2 polypeptide production comprising exposing a transgenic mammal according to Claim 54 to the compound, and measuring CHL2 polypeptide activity or CHL2 polypeptide production in said mammal.

56. An isolated polypeptide comprising the amino acid sequence as set forth in SEQ ID NO: 5 with at least one conservative amino acid substitution selected from the group consisting of: leucine or methionine at position 2; methionine at position 5; lysine at position 6; alanine at position 7; isoleucine at position 8; phenylalanine at position 14; leucine at position 15; threonine at position 23; leucine at position 25; valine at position 27; glutamic acid at position 30; tyrosine at position 32; methionine at position 34; glutamine at position 36; lysine at position 39; alanine at position 41; threonine at position 45; valine at position 55; valine at position 59; asparagine at position 60; proline at position 66; asparagine at position 68; serine or threonine at position 72; valine at position 74; arginine at position 75; arginine at position 94; asparagine at position 99; serine at position 100; lysine at position 105; valine at position 106; tyrosine at position 113; serine at position 116; serine at position 118; arginine at position 120;

leucine at position 123; alanine at position 125; alanine at position 129; alanine at position 139; threonine at position 142; serine at position 144; asparagine at position 147; valine at position 148; serine at position 149; alanine at position 159; alanine at position 160; alanine at position 161; valine at position 164; valine at position 166; valine at position 173; arginine at position 175; aspartic acid at position 177; alanine at position 190; phenylalanine at position 191; arginine at position 192; leucine at position 194; asparagine at position 196; leucine at position 205; alanine at position 210; alanine at position 212; serine at position 213; alanine at position 216; serine at position 217; alanine at position 218; isoleucine at position 219; alanine at position 222; leucine at position 225; phenylalanine at position 226; leucine at position 230; glutamine or arginine at position 233; glutamine at position 241; leucine at position 242; isoleucine at position 244; glutamine or asparagine at position 245; glutamine at position 249; leucine or valine at position 251; alanine at position 256; asparagine at position 257; serine at position 259; alanine at position 260; glutamine at position 261; phenylalanine at position 265; valine at position 268; leucine at position 269; leucine at position 272; valine at position 275; valine at position 278; glutamic acid at position 284; glutamic acid at position 288; alanine or isoleucine at position 292; serine at position 300; isoleucine at position 306; valine at position 313; serine at position 314; leucine at position 319; glutamine at position 323; threonine at position 324; alanine at position 326; alanine at position 327; serine at position 329; serine at position 331; leucine at position 334; asparagine at position 337; valine at position 339; leucine at position 340; serine at position 342; phenylalanine at position 344; glutamic acid at position 349; isoleucine or valine at position 354; methionine at position 356; valine at position 366; methionine or valine at position 368; isoleucine at position 371; leucine at position 375; leucine at position 376; glutamine at position 377; phenylalanine at position 381; asparagine at position 383; isoleucine at position 384; leucine at position 393; arginine at position 395; valine at position 398; alanine or valine at position 399; tyrosine at position 403; asparagine at position 406; isoleucine at position 409; alanine or valine at position 415; isoleucine at position 417; and

leucine at position 421; wherein the polypeptide has an activity of the polypeptide as set forth in SEQ ID NO: 5.

# FIG. 1A

```
cctgccgagg cgtgcacagc ggcagcgctc aacctccccc gcgccgccac cgagggtctt 60

gtcgccccac cgcgccccag acccgcgccg gaccccgcgc cgccgcgccg ccgccagcca 120

gcgccacagg gacactgcac cccggtgacc gcaccccgca gatcccggtt ctctagctag 180

caccttctcc ctctctgcca tagccttttt cttcatttcc ccaactaatt tctctctctc 240

tctctctctc tctctctctc tctcactcac tctctctctc ttctcctcgt cccctcccca 300

ccgtcctctc atcctcacct tagacctctc ctgtccttgg ctcctcttca tctttgcttt 360

tccgactcct caagcagcgg tcctacttgg tcctctgagg acttacttgt gtccttatct 420

cactttctcc cggctcatcc cggggttgtc tgaccttggg acaaggaagg atg gtt    476
                                                    Met Val
                                                      1

ccc ggg gtg agg atc atc ccc tct ttg ctg gga ctc gtg atg ttc tgg   524
Pro Gly Val Arg Ile Ile Pro Ser Leu Leu Gly Leu Val Met Phe Trp
          5                  10                  15

ctc ccg ttg gac tcg caa gca cta tcc cgc tcg ggc aaa gtc tgc ctt   572
Leu Pro Leu Asp Ser Gln Ala Leu Ser Arg Ser Gly Lys Val Cys Leu
     20                  25                  30

ttc ggt gaa aag ata tat acc ccc ggc cag agc tgg cac ccc tac ttg   620
Phe Gly Glu Lys Ile Tyr Thr Pro Gly Gln Ser Trp His Pro Tyr Leu
 35                  40                  45                  50

gaa cca caa ggc acg ata tac tgc gtg cgc tgt acc tgc tct gag aat   668
Glu Pro Gln Gly Thr Ile Tyr Cys Val Arg Cys Thr Cys Ser Glu Asn
                 55                  60                  65

gga cat gtg aat tgt tac cgc ctc cgc tgc cca ccc ctt cac tgc tca   716
Gly His Val Asn Cys Tyr Arg Leu Arg Cys Pro Pro Leu His Cys Ser
             70                  75                  80

cag cct gtg atg gag cca cag caa tgc tgt ccc agg tgt gtg gat cct   764
Gln Pro Val Met Glu Pro Gln Gln Cys Cys Pro Arg Cys Val Asp Pro
         85                  90                  95

cat gtc ccc tct ggc ctc cga gtt ccc cta aag tcc tgc cag ctc aat   812
His Val Pro Ser Gly Leu Arg Val Pro Leu Lys Ser Cys Gln Leu Asn
    100                 105                 110

gag acc aca tac caa cat gga gag atc ttc agt gcc cag gag ctg ttc   860
Glu Thr Thr Tyr Gln His Gly Glu Ile Phe Ser Ala Gln Glu Leu Phe
115                 120                 125                 130

cct gcc cgc ctg tcc aac cag tgt gtc ctg tgt agc tgt att gaa ggc   908
Pro Ala Arg Leu Ser Asn Gln Cys Val Leu Cys Ser Cys Ile Glu Gly
                135                 140                 145
```

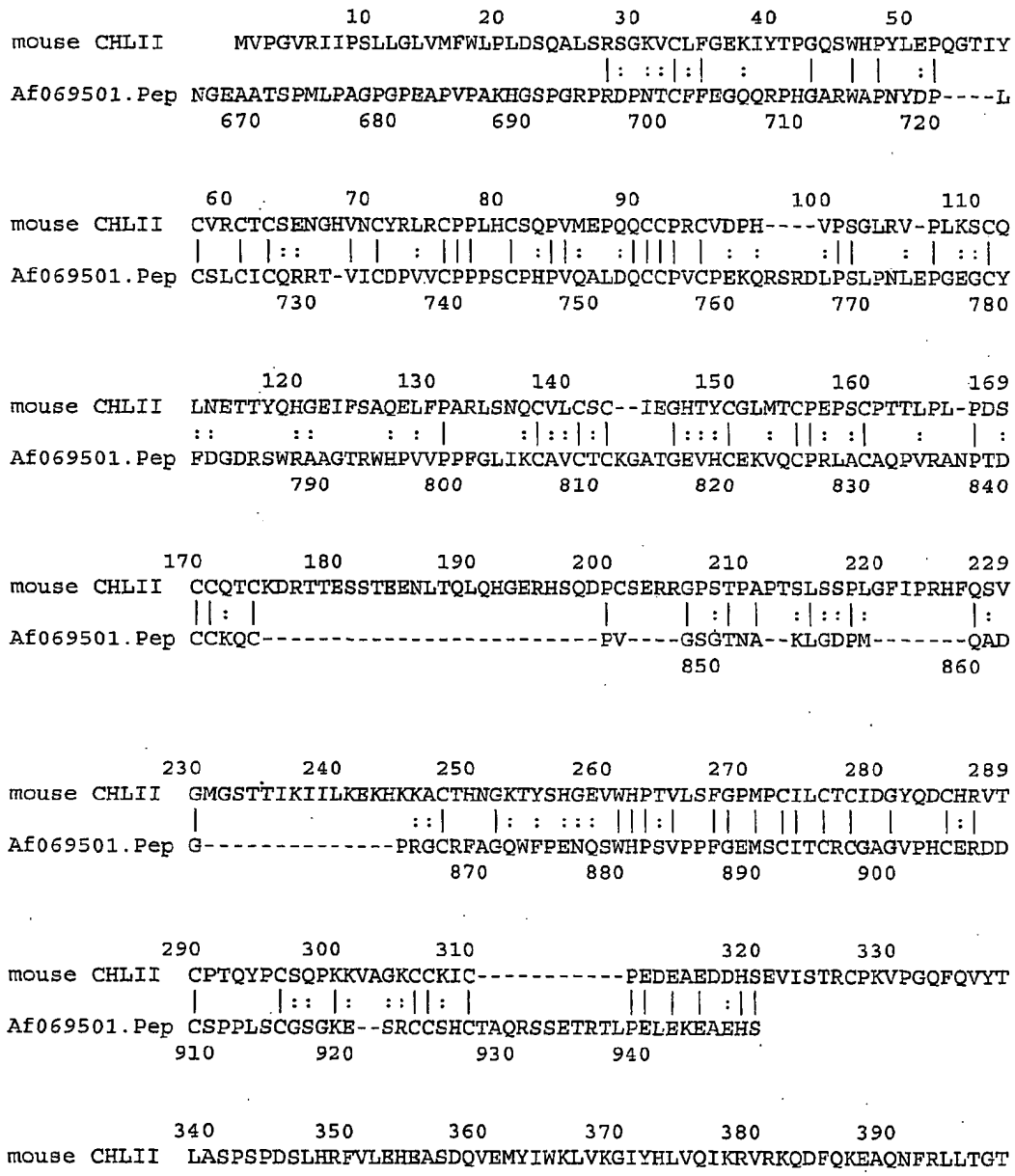
# FIG. 1B

```
cac act tac tgt ggt ctc atg acc tgt cct gaa ccc agc tgc ccc acc   956
His Thr Tyr Cys Gly Leu Met Thr Cys Pro Glu Pro Ser Cys Pro Thr
            150                 155                 160

aca ctc cct ctg cct gat tcc tgc tgt cag acc tgc aaa gac agg aca   1004
Thr Leu Pro Leu Pro Asp Ser Cys Cys Gln Thr Cys Lys Asp Arg Thr
        165                 170                 175

act gag agt tcc aca gaa gaa aac ttg aca cag ctg cag cat gga gag   1052
Thr Glu Ser Ser Thr Glu Glu Asn Leu Thr Gln Leu Gln His Gly Glu
    180                 185                 190

aga cat tcc cag gat cca tgc tcg gag agg aga ggc ccc agc acg cca   1100
Arg His Ser Gln Asp Pro Cys Ser Glu Arg Arg Gly Pro Ser Thr Pro
195                 200                 205                 210

gcc ccc acc agc ctc agc tcc cct ctg ggc ttc atc cct cgc cac ttc   1148
Ala Pro Thr Ser Leu Ser Ser Pro Leu Gly Phe Ile Pro Arg His Phe
                215                 220                 225

cag tca gta gga atg ggc agc aca acc atc aag att atc ttg aag gag   1196
Gln Ser Val Gly Met Gly Ser Thr Thr Ile Lys Ile Ile Leu Lys Glu
            230                 235                 240

aaa cat aaa aaa gct tgc aca cac aat ggg aag aca tac tcc cat ggg   1244
Lys His Lys Lys Ala Cys Thr His Asn Gly Lys Thr Tyr Ser His Gly
        245                 250                 255

gag gtg tgg cac ccc act gtg ctc tcc ttt ggc ccc atg ccc tgc atc   1292
Glu Val Trp His Pro Thr Val Leu Ser Phe Gly Pro Met Pro Cys Ile
    260                 265                 270

ctg tgc aca tgt att gat ggc tac cag gac tgc cac cgt gtg acc tgc   1340
Leu Cys Thr Cys Ile Asp Gly Tyr Gln Asp Cys His Arg Val Thr Cys
275                 280                 285                 290

ccc acc caa tat ccc tgc agt caa ccc aag aaa gtg gct ggg aag tgc   1388
Pro Thr Gln Tyr Pro Cys Ser Gln Pro Lys Lys Val Ala Gly Lys Cys
                295                 300                 305

tgc aag atc tgc cca gag gac gag gcg gaa gat gac cac agt gag gtc   1436
Cys Lys Ile Cys Pro Glu Asp Glu Ala Glu Asp Asp His Ser Glu Val
            310                 315                 320

att tcc acc cgg tgt ccc aag gta cca ggc cag ttc cag gtg tac acg   1484
Ile Ser Thr Arg Cys Pro Lys Val Pro Gly Gln Phe Gln Val Tyr Thr
        325                 330                 335

ttg gca tct cca agc cca gac agc cta cac cgc ttt gtc ctg gag cat   1532
Leu Ala Ser Pro Ser Pro Asp Ser Leu His Arg Phe Val Leu Glu His
    340                 345                 350
```

# 1C

```
gaa gcc tct gac cag gta gag atg tac att tgg aag ctg gtg aaa gga   1580
Glu Ala Ser Asp Gln Val Glu Met Tyr Ile Trp Lys Leu Val Lys Gly
355                 360                 365                 370

atc tac cac ttg gtt cag atc aag aga gtc agg aag caa gat ttc cag   1628
Ile Tyr His Leu Val Gln Ile Lys Arg Val Arg Lys Gln Asp Phe Gln
                375                 380                 385

aaa gag gct cag aac ttc cgg ctg ctc acc ggc acc cat gaa ggt tac   1676
Lys Glu Ala Gln Asn Phe Arg Leu Leu Thr Gly Thr His Glu Gly Tyr
            390                 395                 400

tgg acc gtc ttc cta gcc cag act cca gag ctg aaa gtt aca gcc agc   1724
Trp Thr Val Phe Leu Ala Gln Thr Pro Glu Leu Lys Val Thr Ala Ser
        405                 410                 415

cca gac aaa gtg acc aag aca tta tagcaaggac ctaaagagtt gcagatacga  1778
Pro Asp Lys Val Thr Lys Thr Leu
    420                 425

gttttattgg ttttgttatt atatattaat aaagaagtcg cattaccctc tccccccac 1838

t                                                                 1839
```

# FIG. 2




SCORES Init1: 113 Initn: 313 Opt: 150
Smith-Waterman score: 340; 24.4% identity in 312 aa overlap

# FIG. 3A

```
agacctccct tcctgccctc ctttcctgcc caccgctgct tcctggccct tctccgaccc 60

cgctctagca gcagacctcc tggggtctgt gggttgatct gtggcccctg tgcctccgtg 120

tccttttcgt ctcccttcct cccgactccg ctcccggacc agcggcctga ccctggggaa 180

agg atg gtt ccc gag gtg agg gtc ctc tcc tcc ttg ctg gga ctc gcg   228
    Met Val Pro Glu Val Arg Val Leu Ser Ser Leu Leu Gly Leu Ala
      1               5                  10                  15

ctg ctc tgg ttc ccc ctg gac tcc cac gct cga gcc cgc cca gac atg   276
Leu Leu Trp Phe Pro Leu Asp Ser His Ala Arg Ala Arg Pro Asp Met
                  20                  25                  30

ttc tgc ctt ttc cat ggg aag aga tac tcc ccc ggc gag agc tgg cac   324
Phe Cys Leu Phe His Gly Lys Arg Tyr Ser Pro Gly Glu Ser Trp His
             35                  40                  45

ccc tac ttg gag cca caa ggc ctg atg tac tgc ctg cgc tgt acc tgc   372
Pro Tyr Leu Glu Pro Gln Gly Leu Met Tyr Cys Leu Arg Cys Thr Cys
         50                  55                  60

tca gag ggc gcc cat gtg agt tgt tac cgc ctc cac tgt ccg cct gtc   420
Ser Glu Gly Ala His Val Ser Cys Tyr Arg Leu His Cys Pro Pro Val
     65                  70                  75

cac tgc ccc cag cct gtg acg gag cca cag caa tgc tgt ccc aag tgt   468
His Cys Pro Gln Pro Val Thr Glu Pro Gln Gln Cys Cys Pro Lys Cys
 80                  85                  90                  95

gtg gaa cct cac act ccc tct gga ctc cgg gcc cca cca aag tcc tgc   516
Val Glu Pro His Thr Pro Ser Gly Leu Arg Ala Pro Pro Lys Ser Cys
                100                 105                 110

cag cac aac ggg acc atg tac caa cac gga gag atc ttc agt gcc cat   564
Gln His Asn Gly Thr Met Tyr Gln His Gly Glu Ile Phe Ser Ala His
            115                 120                 125

gag ctg ttc ccc tcc cgc ctg ccc aac cag tgt gtc ctc tgc agc tgc   612
Glu Leu Phe Pro Ser Arg Leu Pro Asn Gln Cys Val Leu Cys Ser Cys
        130                 135                 140

aca gag ggc cag atc tac tgc ggc ctc aca acc tgc ccc gaa cca ggc   660
Thr Glu Gly Gln Ile Tyr Cys Gly Leu Thr Thr Cys Pro Glu Pro Gly
    145                 150                 155

tgc cca gca ccc ctc ccg ctg cca gac tcc tgc tgc caa gcc tgc aaa   708
Cys Pro Ala Pro Leu Pro Leu Pro Asp Ser Cys Cys Gln Ala Cys Lys
160                 165                 170                 175

gat gag gca agt gag caa tcg gat gaa gag gac agt gtg cag tcg ctc   756
Asp Glu Ala Ser Glu Gln Ser Asp Glu Glu Asp Ser Val Gln Ser Leu
                180                 185                 190
```

# FIG. 3B

```
cat ggg gtg aga cat cct cag gat cca tgt tcc agt gat gct ggg aga   804
His Gly Val Arg His Pro Gln Asp Pro Cys Ser Ser Asp Ala Gly Arg
            195                 200                 205

aag aga ggc ccg ggc acc cca gcc ccc act ggc ctc agc gcc cct ctg   852
Lys Arg Gly Pro Gly Thr Pro Ala Pro Thr Gly Leu Ser Ala Pro Leu
        210                 215                 220

agc ttc atc cct cgc cac ttc aga ccc aag gga gca ggc agc aca act   900
Ser Phe Ile Pro Arg His Phe Arg Pro Lys Gly Ala Gly Ser Thr Thr
    225                 230                 235

gtc aag atc gtc ctg aag gag aaa cat aag aaa gcc tgt gtg cat ggc   948
Val Lys Ile Val Leu Lys Glu Lys His Lys Lys Ala Cys Val His Gly
240                 245                 250                 255

ggg aag acg tac tcc cac ggg gag gtg tgg cac ccg gcc ttc cgt gcc   996
Gly Lys Thr Tyr Ser His Gly Glu Val Trp His Pro Ala Phe Arg Ala
                260                 265                 270

ttc ggc ccc ttg ccc tgc atc cta tgc acc tgt gag gat ggc cgc cag   1044
Phe Gly Pro Leu Pro Cys Ile Leu Cys Thr Cys Glu Asp Gly Arg Gln
            275                 280                 285

gac tgc cag cgt gtg acc tgt ccc acc gag tac ccc tgc cgt cac ccc   1092
Asp Cys Gln Arg Val Thr Cys Pro Thr Glu Tyr Pro Cys Arg His Pro
        290                 295                 300

gag aaa gtg gct ggg aag tgc tgc aag att tgc cca gag gac aaa gca   1140
Glu Lys Val Ala Gly Lys Cys Cys Lys Ile Cys Pro Glu Asp Lys Ala
    305                 310                 315

gac cct ggc cac agt gag atc agt tct acc agg tgt ccc aag gca ccg   1188
Asp Pro Gly His Ser Glu Ile Ser Ser Thr Arg Cys Pro Lys Ala Pro
320                 325                 330                 335

ggc cgg gtc ctc gtc cac aca tcg gta tcc cca agc cca gac aac ctg   1236
Gly Arg Val Leu Val His Thr Ser Val Ser Pro Ser Pro Asp Asn Leu
                340                 345                 350

cgt cgc ttt gcc ctg gaa cac gag gcc tcg gac ttg gtg gag atc tac   1284
Arg Arg Phe Ala Leu Glu His Glu Ala Ser Asp Leu Val Glu Ile Tyr
            355                 360                 365

ctc tgg aag ctg gtg aaa gga atc ttc cac ttg act cag atc aag aaa   1332
Leu Trp Lys Leu Val Lys Gly Ile Phe His Leu Thr Gln Ile Lys Lys
        370                 375                 380

gtc agg aag caa gac ttc cag aaa gag gca cag cac ttc cga ctg ctc   1380
Val Arg Lys Gln Asp Phe Gln Lys Glu Ala Gln His Phe Arg Leu Leu
    385                 390                 395
```

FIG. 3C

```
gct ggc ccc cac gaa ggt cac tgg aac gtc ttc cta gcc cag acc ctg   1428
Ala Gly Pro His Glu Gly His Trp Asn Val Phe Leu Ala Gln Thr Leu
400                 405                 410                 415

gag ctg aag gtc acg gcc agt cca gac aaa gtg acc aag aca           1470
Glu Leu Lys Val Thr Ala Ser Pro Asp Lys Val Thr Lys Thr
                420                 425

taacaaagac ctaacagttg cagatatgag ctgtataatt gttgttatta tatattaata 1530

aataagaagt tgcattaccc tcaaaaaaaa aaaaaaaaaa                      1570
```

FIG. 4A

```
          1                                                  50
huCHD     MPSLPAPPAPLLLLGLLLLGSRPARGAGPEPPVLPIRSEKEPLPVRGAAG
huCHL     ---MGGMKYIFSLLFFLLLEG----------------GKTEQVKHSETY
huCHL2    --MVPEVRVLSSLLGLALLWFP---------------LDSHARARPDMF

          51                                                 100
huCHD     CTFGGKVYALDETWHPDLGEPFGVMRCVLCACEAPQWGRRTRGPGRVSCK
huCHL     CMFQDKKYRVGERWHPYL-EPYGLVYCVNCICSENGNVLCSR--------
huCHL2    CLFHGKRYSPGESWHPYL-EPQGLMYCLRCTCSEGAHVSCYR--------

          101                                                150
huCHD     NIKPECPTPACGQPRQLPGHCCQTCPQERSSSERQPSGLSFEYPRDPEHR
huCHL     ---VRCPNVHCLSPVHIPHLCCPRCFEDSLPPVN----NKVTSKSCEYNG
huCHL2    ---LHCPPVHCPQPVTEPQQCCPKCVEPHTPSG-----LRAPPKSCQHNG

          151                                                200
huCHD     SYSDRGEPGAEERARGDGHTDFVALLTGPRSQAVARARVSLLRSSLRFSI
huCHL     TTYQHGELFVAEGLFQNRQPNQCTQCSCSEG--------------------
huCHL2    TMYQHGEIFSAHELFPSRLPNQCVLCSCTEG--------------------

          201                                                250
huCHD     SYRRLDRPTRIRFSDSNGSVLFEHPAAPTQDGLVCGVWRAVPRLSLRLLR
huCHL     ------------NVYCGLKTCPKLTCAFPVSVPDSCCRVCRGDG-ELSW
huCHL2    ------------QIYCGLTTCPEPGCPAPLPLPDSCCQACKDEASEQSD

          251                                                300
huCHD     AEQLHVALVTLTHPSGEVWGPLIRHRALAAETFSAILTLEGPPQQGVGGI
huCHL     EHSDGDIFRQPANREARHSYHRSHYDPPPSRQAGGLSRFP----------
huCHL2    EEDSVQSLHGVRHPQDPCSSDAGRKRGPGTPAPTGLS------------

          301                                                350
huCHD     TLLTLSDTEDSLHFLLLFRGLLEPRSGGLTQVPLRLQILHQGQLLRELQA
huCHL     ------------GARSHRGALMDSQQASGTLVQIVTNNKHKHGQVCVSNG
huCHL2    ------------APLSFIPRHFRPKGAGSTTVKIVLKEKHKK--ACVHGG

          351                                                400
huCHD     NVSAQEPGFAEVLPNLTVQEMDWLVLGELQMALEWAGRPGLRISGHIAAR
huCHL     KTYSHGESWHPNLRAFGIVECVLCTCNVTKQECKKICHCPNRYPCKYPQKI
huCHL2    KTYSHGEVWHPAFRAFGPLPCILCTCEDGRQDCQRVTCPTEYPCRHPEKV

          401                                                450
huCHD     KSCDVLQSVLCGADALIPVQTGAAGSASLTLLGNGSIIYQVQVVGTSSEV
huCHL     DGKCCKVCPGKKAKEELPGQSFDNKGYFCGEETMPVYESVFMEDGETTRK
huCHL2    AGKCCKICP---EDKADPGHSEISSTRCPKAPGRVLVHTSVSPSPDNLRR

          451                                                500
huCHD     VAMTLETKPQRRDQRTVLCHMAGLQPGGHTAVGICPGLGARGAHMLLQNE
huCHL     IALETERPPQ------VEVHVWTIRKG------------------ILQHF
huCHL2    FALEHEASDL------VEIYLWKLVKG------------------IFHLT
```

# FIG. 4B

```
          501                                                 550
huCHD     LFLNVGTKDFPDGELRGHVAALPYCGHSARHDTLPVPLAGALVLPPVKSQ
huCHL     HIEKISKR-MFEELPHFKLVTRTTLSQWKIFTEGEAQISQMCSSRVCRTE
huCHL2    QIKKVRKQDFQKEAQHFRLLAGPHEGHWNVFLAQTLELKVTASPDKVTKT

          551                                                 600
huCHD     AAGHAWLSLDTHCHLHYEVLLAGLGGSEQGTVTAHLLGPPGTPGPRRLLK
huCHL     LEDLVKVLYLERSEKGHC----------------------------------
huCHL2    --------------------------------------------------

          601                                                 650
huCHD     GFYGSEAQGVVKDLEPELLRHLAKGMASLLITTKGSPRGELRGQVHIANQ
huCHL     --------------------------------------------------
huCHL2    --------------------------------------------------

          651                                                 700
huCHD     CEVGGLRLEAAGAEGVRALGAPDTASAAPPVVPGLPALAPAKPGGPGRPR
huCHL     --------------------------------------------------
huCHL2    --------------------------------------------------

          701                                                 750
huCHD     DPNTCFFEGQQRPHGARWAPNYDPLCSLCTCQRRTVICDPVVCPPPSCPH
huCHL     --------------------------------------------------
huCHL2    --------------------------------------------------

          751                                                 800
huCHD     PVQAPDQCCPVCPEKQDVRDLPGLPRSRDPGEGCYFDGDRSWRAAGTRWH
huCHL     --------------------------------------------------
huCHL2    --------------------------------------------------

          801                                                 850
huCHD     PVVPPFGLIKCAVCTCKGGTGEVHCEKVQCPRLACAQPVRVNPTDCCKQC
huCHL     --------------------------------------------------
huCHL2    --------------------------------------------------

          851                                                 900
huCHD     PVGSGAHPQLGDPMQADGPRGCRFAGQWFPESQSWHPSVPPFGEMSCITC
huCHL     --------------------------------------------------
huCHL2    --------------------------------------------------

          901                                                 950
huCHD     RCGAGVPHCERDDCSLPLSCGSGKESRCCSRCTAHRRPAPETRTDPELEK
huCHL     --------------------------------------------------
huCHL2    --------------------------------------------------

          951
huCHD     EAEGS
huCHL     -----
huCHL2    -----
```

FIG. 5

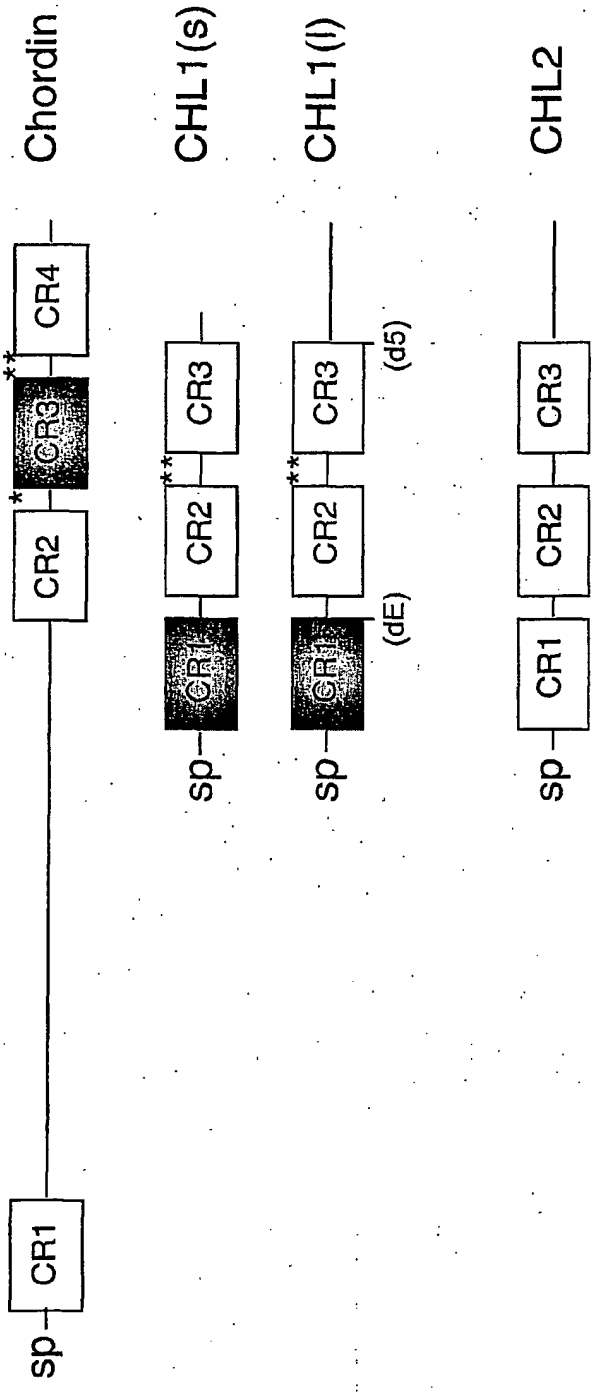

Chordin
CR4
CR3
CR2
CR1
sp
CHL1(s)
CR3
CR2
CR1
sp
CHL1(l)
CR3
(d5)
CR2
(dE)
CR1
sp
CHL2
CR3
CR2
CR1
sp

FIG. 6

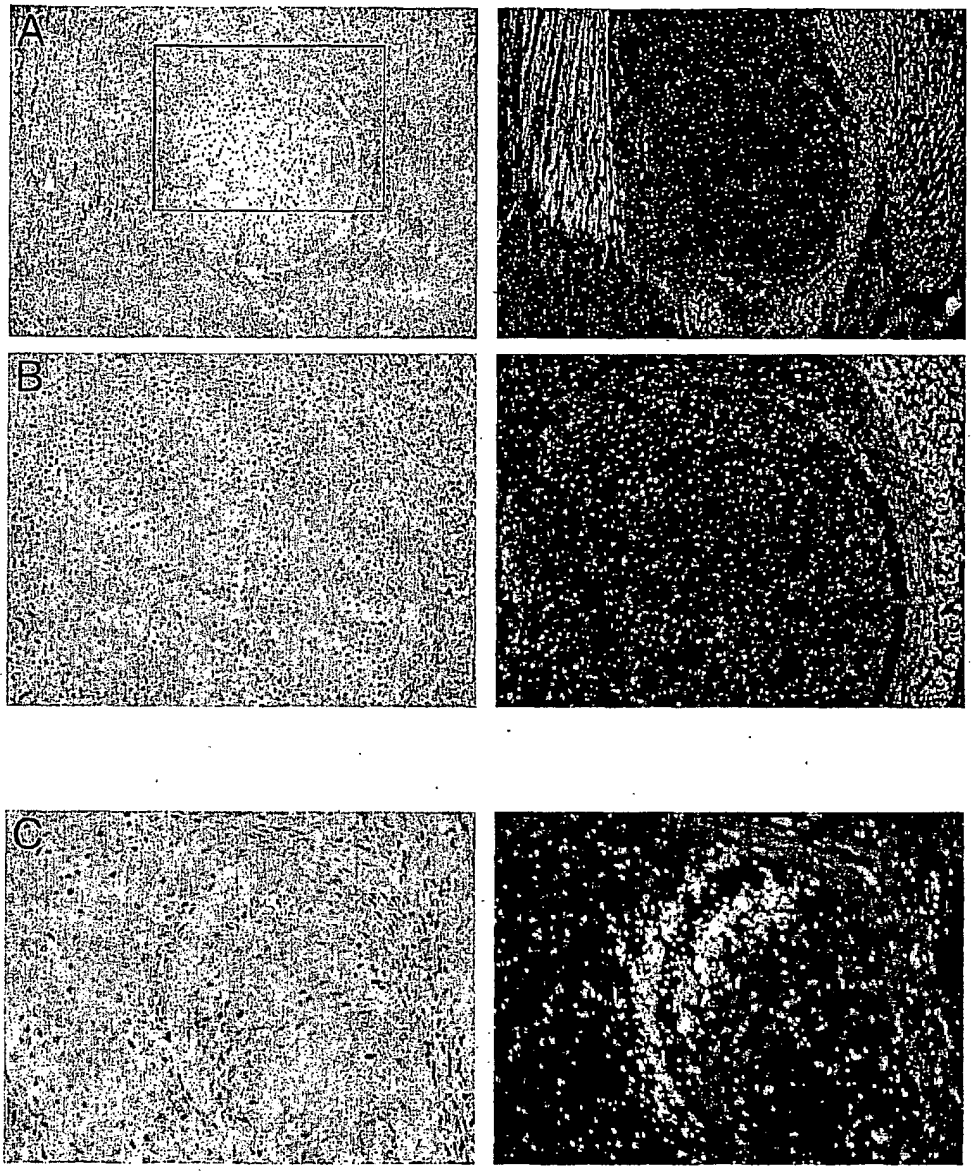


FIG. 7

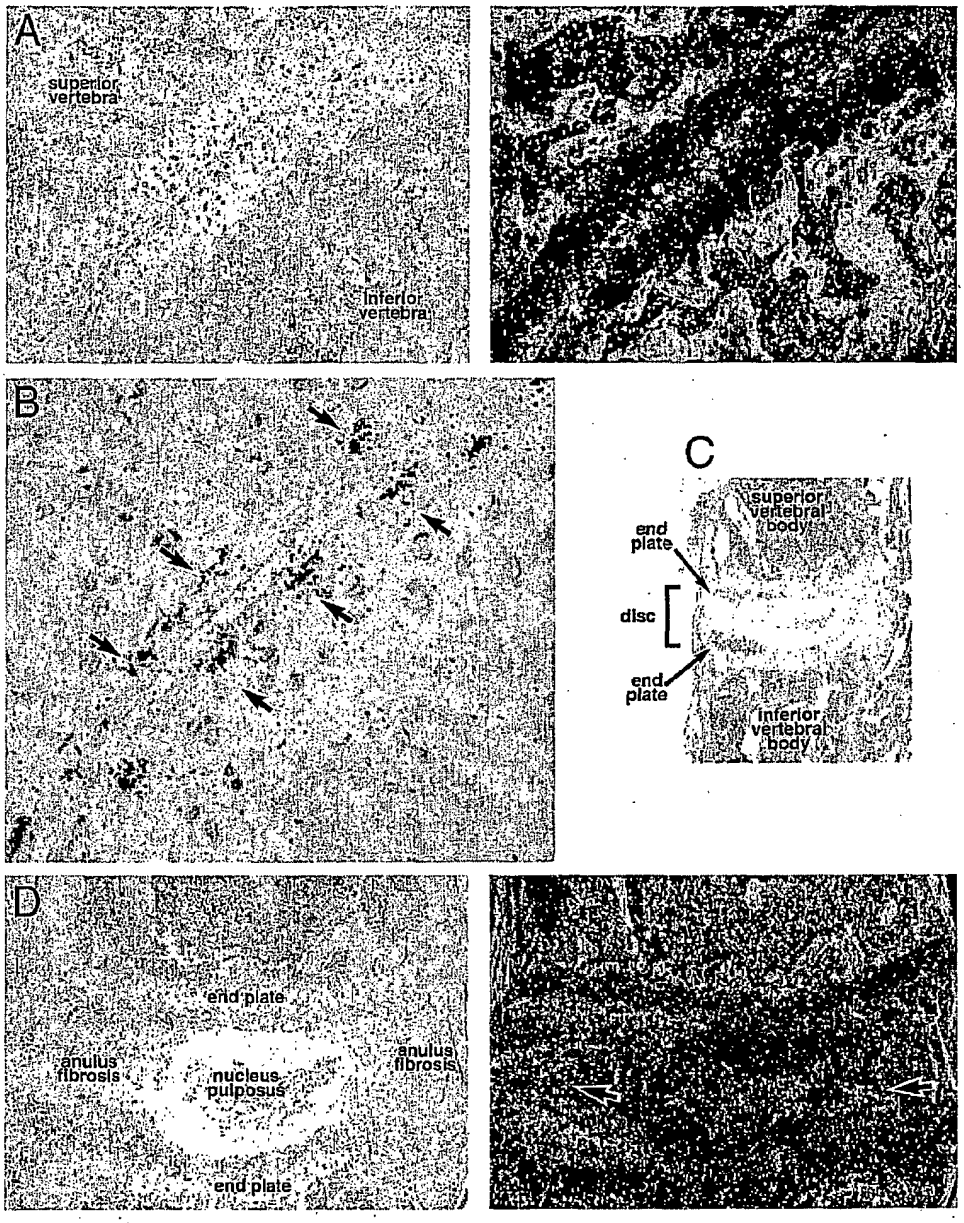

A
superior vertebra
inferior vertebra
B
C
superior vertebral body
end plate
disc
end plate
inferior vertebral body
D
end plate
anulus fibrosis
nucleus pulposus
anulus fibrosis
end plate

FIG. 8

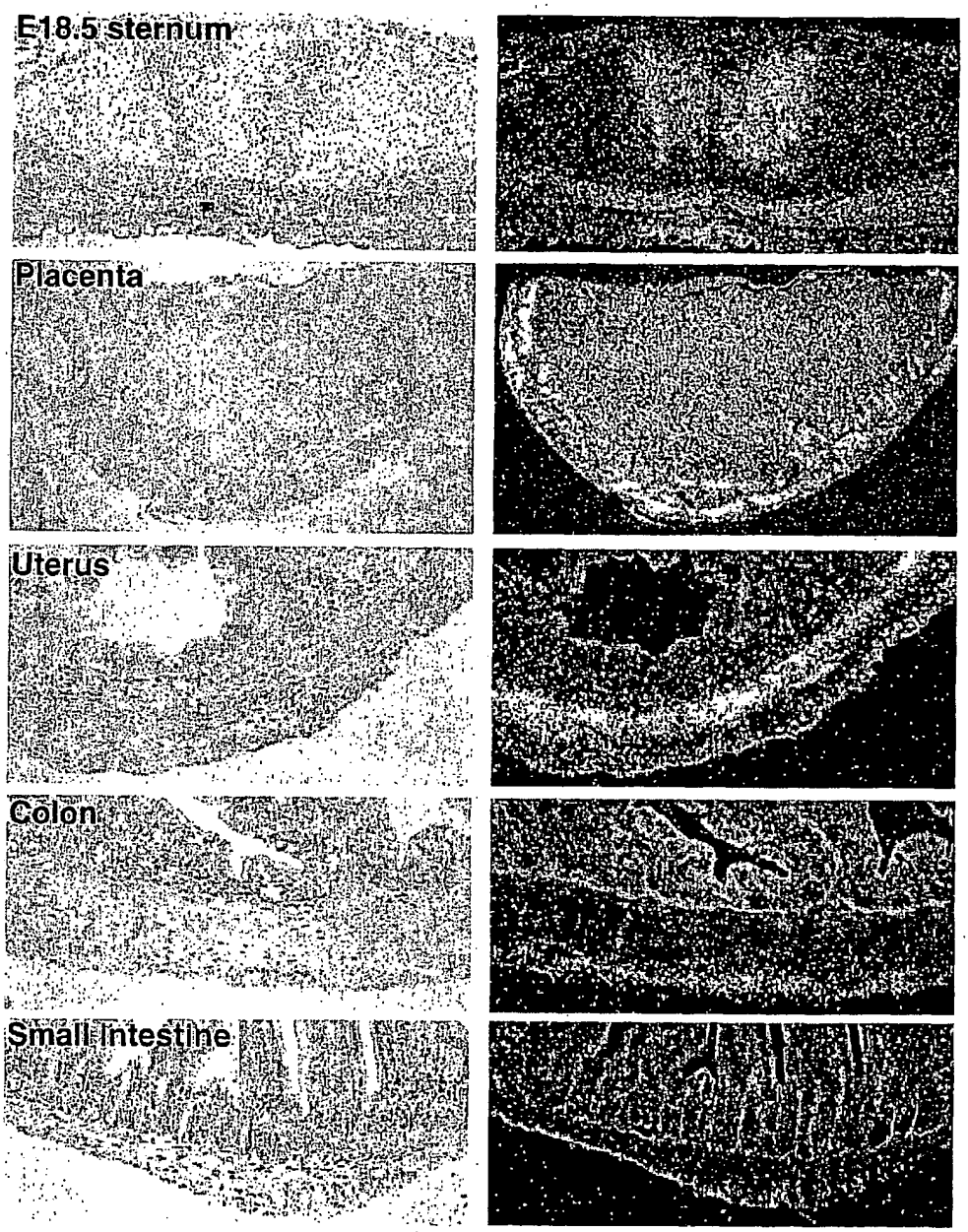
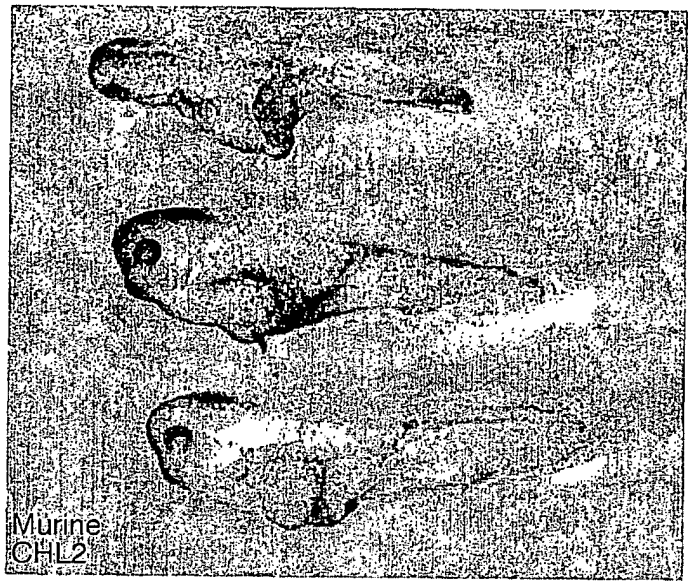


FIG. 9



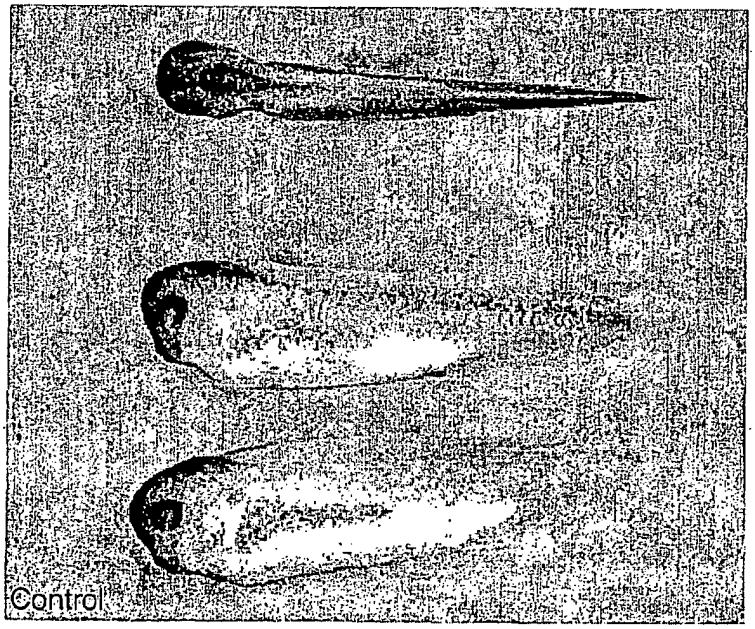



FIG. 10

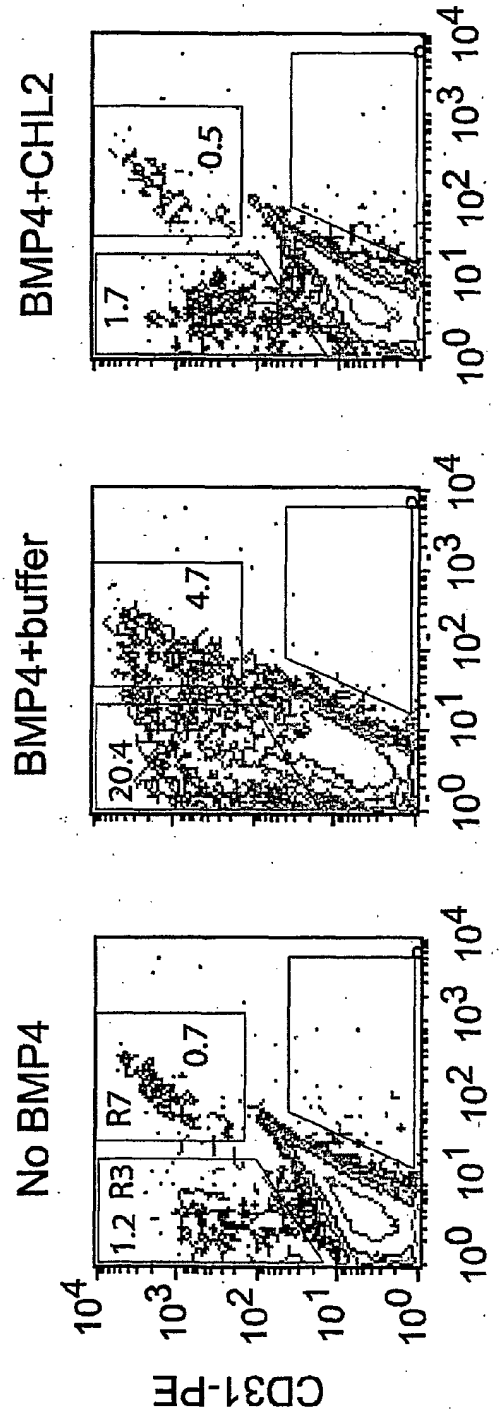

No BMP4
BMP4+buffer
BMP4+CHL2
CD31-PE
CD34-FITC
R3
R7
1.2
0.7
20.4
4.7
1.7
0.5
$10^0$ $10^1$ $10^2$ $10^3$ $10^4$

FIG. 11

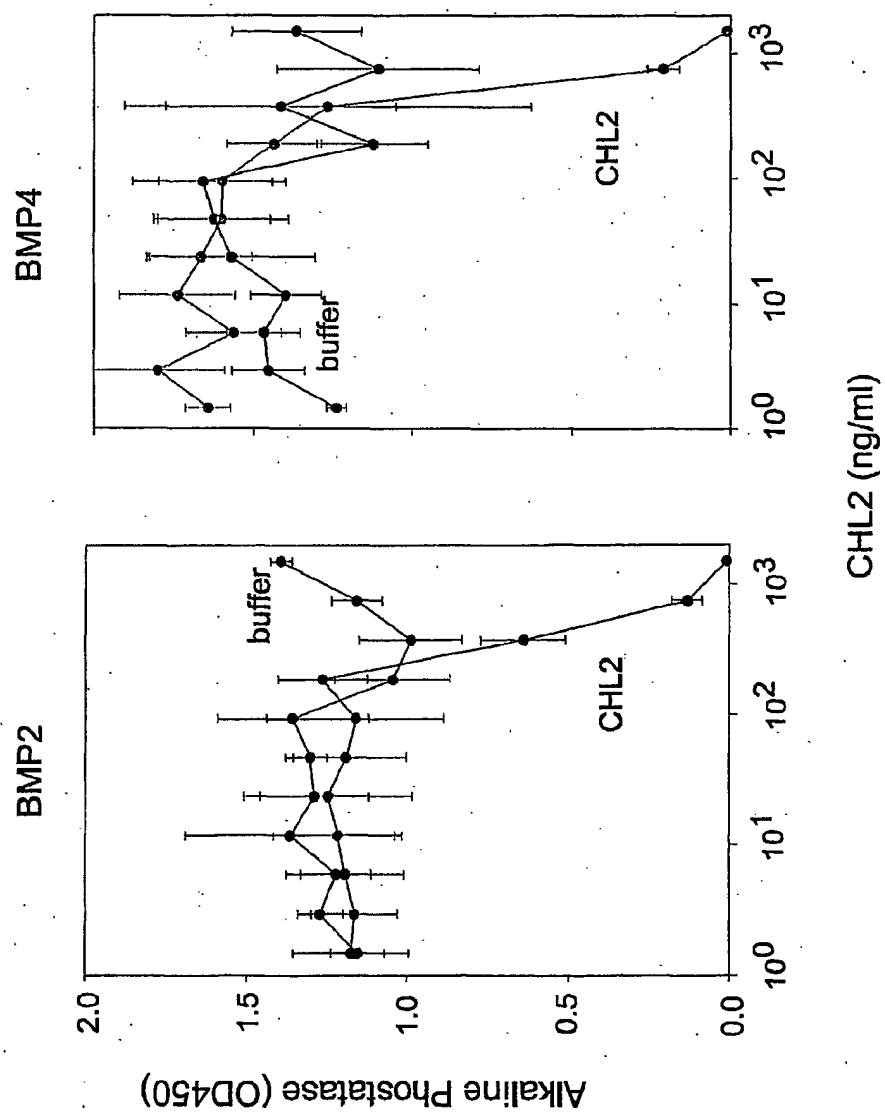
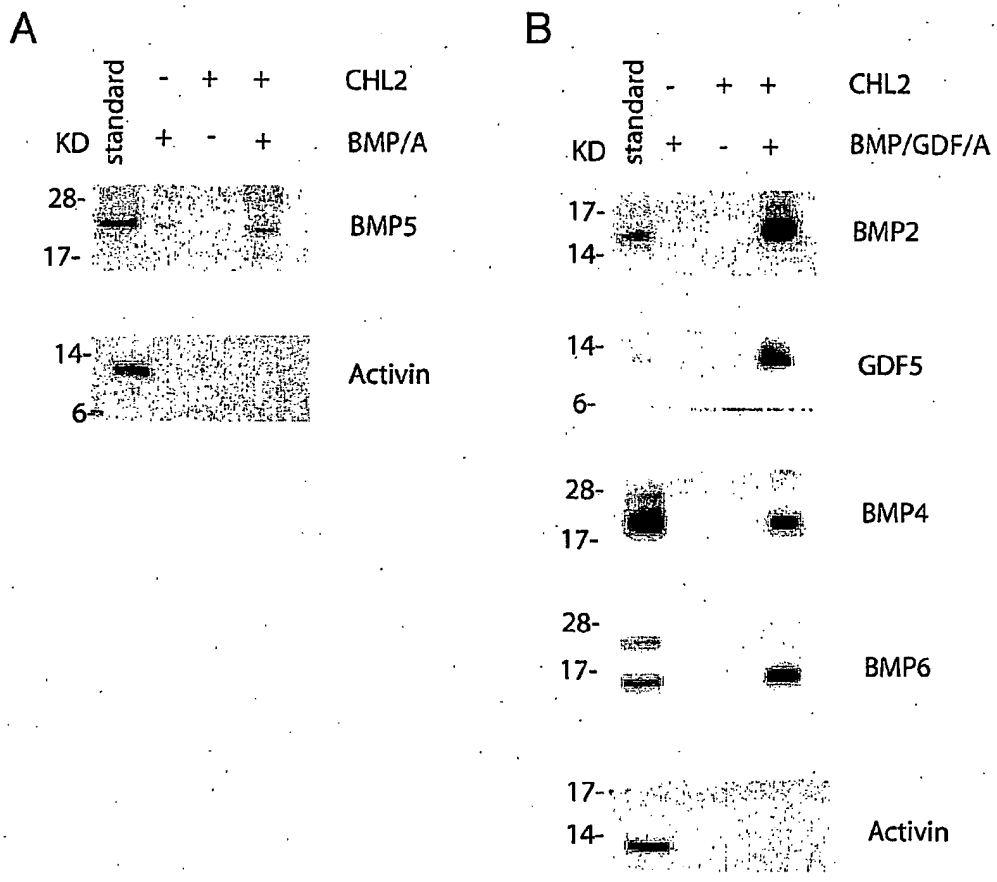

BMP2
BMP4
buffer
CHL2
Alkaline Phosphatase (OD450)
CHL2 (ng/ml)
0.0
0.5
1.0
1.5
2.0
$10^0$
$10^1$
$10^2$
$10^3$

FIG. 12


A
standard
- + +
CHL2
KD
+ - +
BMP/A
28-
17-
BMP5
14-
6-
Activin
B
standard
- + +
CHL2
KD
+ - +
BMP/GDF/A
17-
14-
BMP2
14-
6-
GDF5
28-
17-
BMP4
28-
17-
BMP6
17-
14-
Activin

SEQUENCE LISTING

<110> Zhang, Ke
Cam, Linh
Nakayama, Naoki

<120> Chordin-Like-2 Molecules and Uses Thereof

<130> 01-005

<140> 60/186,462
<141> 2000-03-02

<160> 21

<170> PatentIn Ver. 2.0

<210> 1
<211> 1839
<212> DNA
<213> Mus musculus

<220>
<221> CDS
<222> (471)..(1748)

<220>
<221> sig_peptide
<222> (471)..(530)

<400> 1
cctgccgagg cgtgcacagc ggcagcgctc aacctccccc gcgccgccac cgagggtctt 60

gtcgccccac cgcgccccag acccgcgccg gaccccgcgc cgccgcgccg ccgccagcca 120

gcgccacagg gacactgcac cccggtgacc gcaccccgca gatcccggtt ctctagctag 180

caccttctcc ctctctgcca tagccttttt cttcatttcc ccaactaatt tctctctctc 240

tctctctctc tctctctctc tctcactcac tctctctctc ttctcctcgt cccctcccca 300

ccgtcctctc atcctcacct tagacctctc ctgtccttgg ctcctcttca tctttgcttt 360

tccgactcct caagcagcgg tcctacttgg tcctctgagg acttacttgt gtccttatct 420

cactttctcc cggctcatcc cggggttgtc tgaccttggg acaaggaagg atg gtt 476
Met Val
1

ccc ggg gtg agg atc atc ccc tct ttg ctg gga ctc gtg atg ttc tgg 524
Pro Gly Val Arg Ile Ile Pro Ser Leu Leu Gly Leu Val Met Phe Trp
5 10 15

ctc ccg ttg gac tcg caa gca cta tcc cgc tcg ggc aaa gtc tgc ctt 572
Leu Pro Leu Asp Ser Gln Ala Leu Ser Arg Ser Gly Lys Val Cys Leu
20 25 30

ttc ggt gaa aag ata tat acc ccc ggc cag agc tgg cac ccc tac ttg 620
Phe Gly Glu Lys Ile Tyr Thr Pro Gly Gln Ser Trp His Pro Tyr Leu
35 40 45 50

```
gaa cca caa ggc acg ata tac tgc gtg cgc tgt acc tgc tct gag aat   668
Glu Pro Gln Gly Thr Ile Tyr Cys Val Arg Cys Thr Cys Ser Glu Asn
                 55                  60                  65

gga cat gtg aat tgt tac cgc ctc cgc tgc cca ccc ctt cac tgc tca   716
Gly His Val Asn Cys Tyr Arg Leu Arg Cys Pro Pro Leu His Cys Ser
            70                  75                  80

cag cct gtg atg gag cca cag caa tgc tgt ccc agg tgt gtg gat cct   764
Gln Pro Val Met Glu Pro Gln Gln Cys Cys Pro Arg Cys Val Asp Pro
        85                  90                  95

cat gtc ccc tct ggc ctc cga gtt ccc cta aag tcc tgc cag ctc aat   812
His Val Pro Ser Gly Leu Arg Val Pro Leu Lys Ser Cys Gln Leu Asn
    100                 105                 110

gag acc aca tac caa cat gga gag atc ttc agt gcc cag gag ctg ttc   860
Glu Thr Thr Tyr Gln His Gly Glu Ile Phe Ser Ala Gln Glu Leu Phe
115                 120                 125                 130

cct gcc cgc ctg tcc aac cag tgt gtc ctg tgt agc tgt att gaa ggc   908
Pro Ala Arg Leu Ser Asn Gln Cys Val Leu Cys Ser Cys Ile Glu Gly
                135                 140                 145

cac act tac tgt ggt ctc atg acc tgt cct gaa ccc agc tgc ccc acc   956
His Thr Tyr Cys Gly Leu Met Thr Cys Pro Glu Pro Ser Cys Pro Thr
            150                 155                 160

aca ctc cct ctg cct gat tcc tgc tgt cag acc tgc aaa gac agg aca   1004
Thr Leu Pro Leu Pro Asp Ser Cys Cys Gln Thr Cys Lys Asp Arg Thr
        165                 170                 175

act gag agt tcc aca gaa gaa aac ttg aca cag ctg cag cat gga gag   1052
Thr Glu Ser Ser Thr Glu Glu Asn Leu Thr Gln Leu Gln His Gly Glu
    180                 185                 190

aga cat tcc cag gat cca tgc tcg gag agg aga ggc ccc agc acg cca   1100
Arg His Ser Gln Asp Pro Cys Ser Glu Arg Arg Gly Pro Ser Thr Pro
195                 200                 205                 210

gcc ccc acc agc ctc agc tcc cct ctg ggc ttc atc cct cgc cac ttc   1148
Ala Pro Thr Ser Leu Ser Ser Pro Leu Gly Phe Ile Pro Arg His Phe
                215                 220                 225

cag tca gta gga atg ggc agc aca acc atc aag att atc ttg aag gag   1196
Gln Ser Val Gly Met Gly Ser Thr Thr Ile Lys Ile Ile Leu Lys Glu
            230                 235                 240

aaa cat aaa aaa gct tgc aca cac aat ggg aag aca tac tcc cat ggg   1244
Lys His Lys Lys Ala Cys Thr His Asn Gly Lys Thr Tyr Ser His Gly
        245                 250                 255

gag gtg tgg cac ccc act gtg ctc tcc ttt ggc ccc atg ccc tgc atc   1292
Glu Val Trp His Pro Thr Val Leu Ser Phe Gly Pro Met Pro Cys Ile
    260                 265                 270

ctg tgc aca tgt att gat ggc tac cag gac tgc cac cgt gtg acc tgc   1340
Leu Cys Thr Cys Ile Asp Gly Tyr Gln Asp Cys His Arg Val Thr Cys
275                 280                 285                 290

ccc acc caa tat ccc tgc agt caa ccc aag aaa gtg gct ggg aag tgc   1388
```

```
Pro Thr Gln Tyr Pro Cys Ser Gln Pro Lys Lys Val Ala Gly Lys Cys
                295                 300                 305

tgc aag atc tgc cca gag gac gag gcg gaa gat gac cac agt gag gtc   1436
Cys Lys Ile Cys Pro Glu Asp Glu Ala Glu Asp Asp His Ser Glu Val
            310                 315                 320

att tcc acc cgg tgt ccc aag gta cca ggc cag ttc cag gtg tac acg   1484
Ile Ser Thr Arg Cys Pro Lys Val Pro Gly Gln Phe Gln Val Tyr Thr
        325                 330                 335

ttg gca tct cca agc cca gac agc cta cac cgc ttt gtc ctg gag cat   1532
Leu Ala Ser Pro Ser Pro Asp Ser Leu His Arg Phe Val Leu Glu His
    340                 345                 350

gaa gcc tct gac cag gta gag atg tac att tgg aag ctg gtg aaa gga   1580
Glu Ala Ser Asp Gln Val Glu Met Tyr Ile Trp Lys Leu Val Lys Gly
355                 360                 365                 370

atc tac cac ttg gtt cag atc aag aga gtc agg aag caa gat ttc cag   1628
Ile Tyr His Leu Val Gln Ile Lys Arg Val Arg Lys Gln Asp Phe Gln
                375                 380                 385

aaa gag gct cag aac ttc cgg ctg ctc acc ggc acc cat gaa ggt tac   1676
Lys Glu Ala Gln Asn Phe Arg Leu Leu Thr Gly Thr His Glu Gly Tyr
            390                 395                 400

tgg acc gtc ttc cta gcc cag act cca gag ctg aaa gtt aca gcc agc   1724
Trp Thr Val Phe Leu Ala Gln Thr Pro Glu Leu Lys Val Thr Ala Ser
        405                 410                 415

cca gac aaa gtg acc aag aca tta tagcaaggac ctaaagagtt gcagatacga  1778
Pro Asp Lys Val Thr Lys Thr Leu
    420                 425

gttttattgg ttttgttatt atatattaat aaagaagtcg cattaccctc tccccccac  1838

t                                                                 1839
```

<210> 2
<211> 426
<212> PRT
<213> Mus musculus

<400> 2

```
Met Val Pro Gly Val Arg Ile Ile Pro Ser Leu Leu Gly Leu Val Met
  1               5                  10                  15

Phe Trp Leu Pro Leu Asp Ser Gln Ala Leu Ser Arg Ser Gly Lys Val
             20                  25                  30

Cys Leu Phe Gly Glu Lys Ile Tyr Thr Pro Gly Gln Ser Trp His Pro
         35                  40                  45

Tyr Leu Glu Pro Gln Gly Thr Ile Tyr Cys Val Arg Cys Thr Cys Ser
     50                  55                  60

Glu Asn Gly His Val Asn Cys Tyr Arg Leu Arg Cys Pro Pro Leu His
 65                  70                  75                  80
```

```
Cys Ser Gln Pro Val Met Glu Pro Gln Gln Cys Cys Pro Arg Cys Val
                 85                  90                  95

Asp Pro His Val Pro Ser Gly Leu Arg Val Pro Leu Lys Ser Cys Gln
            100                 105                 110

Leu Asn Glu Thr Thr Tyr Gln His Gly Glu Ile Phe Ser Ala Gln Glu
        115                 120                 125

Leu Phe Pro Ala Arg Leu Ser Asn Gln Cys Val Leu Cys Ser Cys Ile
    130                 135                 140

Glu Gly His Thr Tyr Cys Gly Leu Met Thr Cys Pro Glu Pro Ser Cys
145                 150                 155                 160

Pro Thr Thr Leu Pro Leu Pro Asp Ser Cys Cys Gln Thr Cys Lys Asp
                165                 170                 175

Arg Thr Thr Glu Ser Ser Thr Glu Glu Asn Leu Thr Gln Leu Gln His
            180                 185                 190

Gly Glu Arg His Ser Gln Asp Pro Cys Ser Glu Arg Arg Gly Pro Ser
        195                 200                 205

Thr Pro Ala Pro Thr Ser Leu Ser Ser Pro Leu Gly Phe Ile Pro Arg
    210                 215                 220

His Phe Gln Ser Val Gly Met Gly Ser Thr Thr Ile Lys Ile Ile Leu
225                 230                 235                 240

Lys Glu Lys His Lys Lys Ala Cys Thr His Asn Gly Lys Thr Tyr Ser
                245                 250                 255

His Gly Glu Val Trp His Pro Thr Val Leu Ser Phe Gly Pro Met Pro
            260                 265                 270

Cys Ile Leu Cys Thr Cys Ile Asp Gly Tyr Gln Asp Cys His Arg Val
        275                 280                 285

Thr Cys Pro Thr Gln Tyr Pro Cys Ser Gln Pro Lys Lys Val Ala Gly
    290                 295                 300

Lys Cys Cys Lys Ile Cys Pro Glu Asp Glu Ala Glu Asp Asp His Ser
305                 310                 315                 320

Glu Val Ile Ser Thr Arg Cys Pro Lys Val Pro Gly Gln Phe Gln Val
                325                 330                 335

Tyr Thr Leu Ala Ser Pro Ser Pro Asp Ser Leu His Arg Phe Val Leu
            340                 345                 350

Glu His Glu Ala Ser Asp Gln Val Glu Met Tyr Ile Trp Lys Leu Val
        355                 360                 365

Lys Gly Ile Tyr His Leu Val Gln Ile Lys Arg Val Arg Lys Gln Asp
    370                 375                 380

Phe Gln Lys Glu Ala Gln Asn Phe Arg Leu Leu Thr Gly Thr His Glu
385                 390                 395                 400

Gly Tyr Trp Thr Val Phe Leu Ala Gln Thr Pro Glu Leu Lys Val Thr
```

405 410 415

Ala Ser Pro Asp Lys Val Thr Lys Thr Leu
420 425

<210> 3
<211> 405
<212> PRT
<213> Mus musculus

<400> 3
Asp Ser Gln Ala Leu Ser Arg Ser Gly Lys Val Cys Leu Phe Gly Glu
1 5 10 15

Lys Ile Tyr Thr Pro Gly Gln Ser Trp His Pro Tyr Leu Glu Pro Gln
20 25 30

Gly Thr Ile Tyr Cys Val Arg Cys Thr Cys Ser Glu Asn Gly His Val
35 40 45

Asn Cys Tyr Arg Leu Arg Cys Pro Pro Leu His Cys Ser Gln Pro Val
50 55 60

Met Glu Pro Gln Gln Cys Cys Pro Arg Cys Val Asp Pro His Val Pro
65 70 75 80

Ser Gly Leu Arg Val Pro Leu Lys Ser Cys Gln Leu Asn Glu Thr Thr
85 90 95

Tyr Gln His Gly Glu Ile Phe Ser Ala Gln Glu Leu Phe Pro Ala Arg
100 105 110

Leu Ser Asn Gln Cys Val Leu Cys Ser Cys Ile Glu Gly His Thr Tyr
115 120 125

Cys Gly Leu Met Thr Cys Pro Glu Pro Ser Cys Pro Thr Thr Leu Pro
130 135 140

Leu Pro Asp Ser Cys Cys Gln Thr Cys Lys Asp Arg Thr Thr Glu Ser
145 150 155 160

Ser Thr Glu Glu Asn Leu Thr Gln Leu Gln His Gly Glu Arg His Ser
165 170 175

Gln Asp Pro Cys Ser Glu Arg Arg Gly Pro Ser Thr Pro Ala Pro Thr
180 185 190

Ser Leu Ser Ser Pro Leu Gly Phe Ile Pro Arg His Phe Gln Ser Val
195 200 205

Gly Met Gly Ser Thr Thr Ile Lys Ile Ile Leu Lys Glu Lys His Lys
210 215 220

Lys Ala Cys Thr His Asn Gly Lys Thr Tyr Ser His Gly Glu Val Trp
225 230 235 240

His Pro Thr Val Leu Ser Phe Gly Pro Met Pro Cys Ile Leu Cys Thr
245 250 255

Cys Ile Asp Gly Tyr Gln Asp Cys His Arg Val Thr Cys Pro Thr Gln

260 265 270

Tyr Pro Cys Ser Gln Pro Lys Lys Val Ala Gly Lys Cys Cys Lys Ile
275 280 285

Cys Pro Glu Asp Glu Ala Glu Asp Asp His Ser Glu Val Ile Ser Thr
290 295 300

Arg Cys Pro Lys Val Pro Gly Gln Phe Gln Val Tyr Thr Leu Ala Ser
305 310 315 320

Pro Ser Pro Asp Ser Leu His Arg Phe Val Leu Glu His Glu Ala Ser
325 330 335

Asp Gln Val Glu Met Tyr Ile Trp Lys Leu Val Lys Gly Ile Tyr His
340 345 350

Leu Val Gln Ile Lys Arg Val Arg Lys Gln Asp Phe Gln Lys Glu Ala
355 360 365

Gln Asn Phe Arg Leu Leu Thr Gly Thr His Glu Gly Tyr Trp Thr Val
370 375 380

Phe Leu Ala Gln Thr Pro Glu Leu Lys Val Thr Ala Ser Pro Asp Lys
385 390 395 400

Val Thr Lys Thr Leu
405

<210> 4
<211> 1570
<212> DNA
<213> Homo sapiens

<220>
<221> CDS
<222> (184)..(1470)

<220>
<221> sig_peptide
<222> (184)..(243)

<400> 4
agacctccct tcctgccctc ctttcctgcc caccgctgct tcctggccct tctccgaccc 60

cgctctagca gcagacctcc tggggtctgt gggttgatct gtggcccctg tgcctccgtg 120

tccttttcgt ctcccttcct cccgactccg ctcccggacc agcggcctga ccctggggaa 180

agg atg gtt ccc gag gtg agg gtc ctc tcc tcc ttg ctg gga ctc gcg 228
Met Val Pro Glu Val Arg Val Leu Ser Ser Leu Leu Gly Leu Ala
1 5 10 15

ctg ctc tgg ttc ccc ctg gac tcc cac gct cga gcc cgc cca gac atg 276
Leu Leu Trp Phe Pro Leu Asp Ser His Ala Arg Ala Arg Pro Asp Met
20 25 30

ttc tgc ctt ttc cat ggg aag aga tac tcc ccc ggc gag agc tgg cac 324
Phe Cys Leu Phe His Gly Lys Arg Tyr Ser Pro Gly Glu Ser Trp His
35 40 45

```
ccc tac ttg gag cca caa ggc ctg atg tac tgc ctg cgc tgt acc tgc   372
Pro Tyr Leu Glu Pro Gln Gly Leu Met Tyr Cys Leu Arg Cys Thr Cys
         50                  55                  60

tca gag ggc gcc cat gtg agt tgt tac cgc ctc cac tgt ccg cct gtc   420
Ser Glu Gly Ala His Val Ser Cys Tyr Arg Leu His Cys Pro Pro Val
     65                  70                  75

cac tgc ccc cag cct gtg acg gag cca cag caa tgc tgt ccc aag tgt   468
His Cys Pro Gln Pro Val Thr Glu Pro Gln Gln Cys Cys Pro Lys Cys
 80                  85                  90                  95

gtg gaa cct cac act ccc tct gga ctc cgg gcc cca cca aag tcc tgc   516
Val Glu Pro His Thr Pro Ser Gly Leu Arg Ala Pro Pro Lys Ser Cys
                100                 105                 110

cag cac aac ggg acc atg tac caa cac gga gag atc ttc agt gcc cat   564
Gln His Asn Gly Thr Met Tyr Gln His Gly Glu Ile Phe Ser Ala His
            115                 120                 125

gag ctg ttc ccc tcc cgc ctg ccc aac cag tgt gtc ctc tgc agc tgc   612
Glu Leu Phe Pro Ser Arg Leu Pro Asn Gln Cys Val Leu Cys Ser Cys
        130                 135                 140

aca gag ggc cag atc tac tgc ggc ctc aca acc tgc ccc gaa cca ggc   660
Thr Glu Gly Gln Ile Tyr Cys Gly Leu Thr Thr Cys Pro Glu Pro Gly
    145                 150                 155

tgc cca gca ccc ctc ccg ctg cca gac tcc tgc tgc caa gcc tgc aaa   708
Cys Pro Ala Pro Leu Pro Leu Pro Asp Ser Cys Cys Gln Ala Cys Lys
160                 165                 170                 175

gat gag gca agt gag caa tcg gat gaa gag gac agt gtg cag tcg ctc   756
Asp Glu Ala Ser Glu Gln Ser Asp Glu Glu Asp Ser Val Gln Ser Leu
                180                 185                 190

cat ggg gtg aga cat cct cag gat cca tgt tcc agt gat gct ggg aga   804
His Gly Val Arg His Pro Gln Asp Pro Cys Ser Ser Asp Ala Gly Arg
            195                 200                 205

aag aga ggc ccg ggc acc cca gcc ccc act ggc ctc agc gcc cct ctg   852
Lys Arg Gly Pro Gly Thr Pro Ala Pro Thr Gly Leu Ser Ala Pro Leu
        210                 215                 220

agc ttc atc cct cgc cac ttc aga ccc aag gga gca ggc agc aca act   900
Ser Phe Ile Pro Arg His Phe Arg Pro Lys Gly Ala Gly Ser Thr Thr
    225                 230                 235

gtc aag atc gtc ctg aag gag aaa cat aag aaa gcc tgt gtg cat ggc   948
Val Lys Ile Val Leu Lys Glu Lys His Lys Lys Ala Cys Val His Gly
240                 245                 250                 255

ggg aag acg tac tcc cac ggg gag gtg tgg cac ccg gcc ttc cgt gcc   996
Gly Lys Thr Tyr Ser His Gly Glu Val Trp His Pro Ala Phe Arg Ala
                260                 265                 270

ttc ggc ccc ttg ccc tgc atc cta tgc acc tgt gag gat ggc cgc cag   1044
Phe Gly Pro Leu Pro Cys Ile Leu Cys Thr Cys Glu Asp Gly Arg Gln
            275                 280                 285
```

```
gac tgc cag cgt gtg acc tgt ccc acc gag tac ccc tgc cgt cac ccc   1092
Asp Cys Gln Arg Val Thr Cys Pro Thr Glu Tyr Pro Cys Arg His Pro
        290                 295                 300

gag aaa gtg gct ggg aag tgc tgc aag att tgc cca gag gac aaa gca   1140
Glu Lys Val Ala Gly Lys Cys Cys Lys Ile Cys Pro Glu Asp Lys Ala
    305                 310                 315

gac cct ggc cac agt gag atc agt tct acc agg tgt ccc aag gca ccg   1188
Asp Pro Gly His Ser Glu Ile Ser Ser Thr Arg Cys Pro Lys Ala Pro
320                 325                 330                 335

ggc cgg gtc ctc gtc cac aca tcg gta tcc cca agc cca gac aac ctg   1236
Gly Arg Val Leu Val His Thr Ser Val Ser Pro Ser Pro Asp Asn Leu
                340                 345                 350

cgt cgc ttt gcc ctg gaa cac gag gcc tcg gac ttg gtg gag atc tac   1284
Arg Arg Phe Ala Leu Glu His Glu Ala Ser Asp Leu Val Glu Ile Tyr
            355                 360                 365

ctc tgg aag ctg gtg aaa gga atc ttc cac ttg act cag atc aag aaa   1332
Leu Trp Lys Leu Val Lys Gly Ile Phe His Leu Thr Gln Ile Lys Lys
        370                 375                 380

gtc agg aag caa gac ttc cag aaa gag gca cag cac ttc cga ctg ctc   1380
Val Arg Lys Gln Asp Phe Gln Lys Glu Ala Gln His Phe Arg Leu Leu
    385                 390                 395

gct ggc ccc cac gaa ggt cac tgg aac gtc ttc cta gcc cag acc ctg   1428
Ala Gly Pro His Glu Gly His Trp Asn Val Phe Leu Ala Gln Thr Leu
400                 405                 410                 415

gag ctg aag gtc acg gcc agt cca gac aaa gtg acc aag aca           1470
Glu Leu Lys Val Thr Ala Ser Pro Asp Lys Val Thr Lys Thr
                420                 425

taacaaagac ctaacagttg cagatatgag ctgtataatt gttgttatta tatattaata 1530

aataagaagt tgcattaccc tcaaaaaaaa aaaaaaaaaa                       1570


<210> 5
<211> 429
<212> PRT
<213> Homo sapiens

<400> 5
Met Val Pro Glu Val Arg Val Leu Ser Ser Leu Leu Gly Leu Ala Leu
  1               5                  10                  15

Leu Trp Phe Pro Leu Asp Ser His Ala Arg Ala Arg Pro Asp Met Phe
             20                  25                  30

Cys Leu Phe His Gly Lys Arg Tyr Ser Pro Gly Glu Ser Trp His Pro
         35                  40                  45

Tyr Leu Glu Pro Gln Gly Leu Met Tyr Cys Leu Arg Cys Thr Cys Ser
     50                  55                  60

Glu Gly Ala His Val Ser Cys Tyr Arg Leu His Cys Pro Pro Val His
 65                  70                  75                  80
```

```
Cys Pro Gln Pro Val Thr Glu Pro Gln Gln Cys Cys Pro Lys Cys Val
                 85                  90                  95

Glu Pro His Thr Pro Ser Gly Leu Arg Ala Pro Pro Lys Ser Cys Gln
            100                 105                 110

His Asn Gly Thr Met Tyr Gln His Gly Glu Ile Phe Ser Ala His Glu
        115                 120                 125

Leu Phe Pro Ser Arg Leu Pro Asn Gln Cys Val Leu Cys Ser Cys Thr
    130                 135                 140

Glu Gly Gln Ile Tyr Cys Gly Leu Thr Thr Cys Pro Glu Pro Gly Cys
145                 150                 155                 160

Pro Ala Pro Leu Pro Leu Pro Asp Ser Cys Cys Gln Ala Cys Lys Asp
                165                 170                 175

Glu Ala Ser Glu Gln Ser Asp Glu Glu Asp Ser Val Gln Ser Leu His
            180                 185                 190

Gly Val Arg His Pro Gln Asp Pro Cys Ser Ser Asp Ala Gly Arg Lys
        195                 200                 205

Arg Gly Pro Gly Thr Pro Ala Pro Thr Gly Leu Ser Ala Pro Leu Ser
    210                 215                 220

Phe Ile Pro Arg His Phe Arg Pro Lys Gly Ala Gly Ser Thr Thr Val
225                 230                 235                 240

Lys Ile Val Leu Lys Glu Lys His Lys Lys Ala Cys Val His Gly Gly
                245                 250                 255

Lys Thr Tyr Ser His Gly Glu Val Trp His Pro Ala Phe Arg Ala Phe
            260                 265                 270

Gly Pro Leu Pro Cys Ile Leu Cys Thr Cys Glu Asp Gly Arg Gln Asp
        275                 280                 285

Cys Gln Arg Val Thr Cys Pro Thr Glu Tyr Pro Cys Arg His Pro Glu
    290                 295                 300

Lys Val Ala Gly Lys Cys Cys Lys Ile Cys Pro Glu Asp Lys Ala Asp
305                 310                 315                 320

Pro Gly His Ser Glu Ile Ser Ser Thr Arg Cys Pro Lys Ala Pro Gly
                325                 330                 335

Arg Val Leu Val His Thr Ser Val Ser Pro Ser Pro Asp Asn Leu Arg
            340                 345                 350

Arg Phe Ala Leu Glu His Glu Ala Ser Asp Leu Val Glu Ile Tyr Leu
        355                 360                 365

Trp Lys Leu Val Lys Gly Ile Phe His Leu Thr Gln Ile Lys Lys Val
    370                 375                 380

Arg Lys Gln Asp Phe Gln Lys Glu Ala Gln His Phe Arg Leu Leu Ala
385                 390                 395                 400
```

```
Gly Pro His Glu Gly His Trp Asn Val Phe Leu Ala Gln Thr Leu Glu
                405                 410                 415

Leu Lys Val Thr Ala Ser Pro Asp Lys Val Thr Lys Thr
            420                 425
```

<210> 6
<211> 408
<212> PRT
<213> Homo sapiens

<400> 6

```
Asp Ser His Ala Arg Ala Arg Pro Asp Met Phe Cys Leu Phe His Gly
  1               5                  10                  15

Lys Arg Tyr Ser Pro Gly Glu Ser Trp His Pro Tyr Leu Glu Pro Gln
             20                  25                  30

Gly Leu Met Tyr Cys Leu Arg Cys Thr Cys Ser Glu Gly Ala His Val
         35                  40                  45

Ser Cys Tyr Arg Leu His Cys Pro Pro Val His Cys Pro Gln Pro Val
     50                  55                  60

Thr Glu Pro Gln Gln Cys Cys Pro Lys Cys Val Glu Pro His Thr Pro
 65                  70                  75                  80

Ser Gly Leu Arg Ala Pro Pro Lys Ser Cys Gln His Asn Gly Thr Met
                85                  90                  95

Tyr Gln His Gly Glu Ile Phe Ser Ala His Glu Leu Phe Pro Ser Arg
            100                 105                 110

Leu Pro Asn Gln Cys Val Leu Cys Ser Cys Thr Glu Gly Gln Ile Tyr
        115                 120                 125

Cys Gly Leu Thr Thr Cys Pro Glu Pro Gly Cys Pro Ala Pro Leu Pro
    130                 135                 140

Leu Pro Asp Ser Cys Cys Gln Ala Cys Lys Asp Glu Ala Ser Glu Gln
145                 150                 155                 160

Ser Asp Glu Glu Asp Ser Val Gln Ser Leu His Gly Val Arg His Pro
                165                 170                 175

Gln Asp Pro Cys Ser Ser Asp Ala Gly Arg Lys Arg Gly Pro Gly Thr
            180                 185                 190

Pro Ala Pro Thr Gly Leu Ser Ala Pro Leu Ser Phe Ile Pro Arg His
        195                 200                 205

Phe Arg Pro Lys Gly Ala Gly Ser Thr Thr Val Lys Ile Val Leu Lys
    210                 215                 220

Glu Lys His Lys Lys Ala Cys Val His Gly Gly Lys Thr Tyr Ser His
225                 230                 235                 240

Gly Glu Val Trp His Pro Ala Phe Arg Ala Phe Gly Pro Leu Pro Cys
                245                 250                 255
```

```
Ile Leu Cys Thr Cys Glu Asp Gly Arg Gln Asp Cys Gln Arg Val Thr
            260                 265                 270

Cys Pro Thr Glu Tyr Pro Cys Arg His Pro Glu Lys Val Ala Gly Lys
        275                 280                 285

Cys Cys Lys Ile Cys Pro Glu Asp Lys Ala Asp Pro Gly His Ser Glu
    290                 295                 300

Ile Ser Ser Thr Arg Cys Pro Lys Ala Pro Gly Arg Val Leu Val His
305                 310                 315                 320

Thr Ser Val Ser Pro Ser Pro Asp Asn Leu Arg Arg Phe Ala Leu Glu
                325                 330                 335

His Glu Ala Ser Asp Leu Val Glu Ile Tyr Leu Trp Lys Leu Val Lys
            340                 345                 350

Gly Ile Phe His Leu Thr Gln Ile Lys Lys Val Arg Lys Gln Asp Phe
        355                 360                 365

Gln Lys Glu Ala Gln His Phe Arg Leu Leu Ala Gly Pro His Glu Gly
    370                 375                 380

His Trp Asn Val Phe Leu Ala Gln Thr Leu Glu Leu Lys Val Thr Ala
385                 390                 395                 400

Ser Pro Asp Lys Val Thr Lys Thr
                405
```

<210> 7
<211> 283
<212> PRT
<213> Mus musculus

<400> 7

```
Asn Gly Glu Ala Ala Thr Ser Pro Met Leu Pro Ala Gly Pro Gly Pro
  1               5                  10                  15

Glu Ala Pro Val Pro Ala Lys His Gly Ser Pro Gly Arg Pro Arg Asp
             20                  25                  30

Pro Asn Thr Cys Phe Phe Glu Gly Gln Gln Arg Pro His Gly Ala Arg
         35                  40                  45

Trp Ala Pro Asn Tyr Asp Pro Leu Cys Ser Leu Cys Ile Cys Gln Arg
     50                  55                  60

Arg Thr Val Ile Cys Asp Pro Val Val Cys Pro Pro Pro Ser Cys Pro
 65                  70                  75                  80

His Pro Val Gln Ala Leu Asp Gln Cys Cys Pro Val Cys Pro Glu Lys
                 85                  90                  95

Gln Arg Ser Arg Asp Leu Pro Ser Leu Pro Asn Leu Glu Pro Gly Glu
            100                 105                 110

Gly Cys Tyr Phe Asp Gly Asp Arg Ser Trp Arg Ala Ala Gly Thr Arg
        115                 120                 125
```

```
Trp His Pro Val Val Pro Pro Phe Gly Leu Ile Lys Cys Ala Val Cys
    130                 135                 140

Thr Cys Lys Gly Ala Thr Gly Glu Val His Cys Glu Lys Val Gln Cys
145                 150                 155                 160

Pro Arg Leu Ala Cys Ala Gln Pro Val Arg Ala Asn Pro Thr Asp Cys
                165                 170                 175

Cys Lys Gln Cys Pro Val Gly Ser Gly Thr Asn Ala Lys Leu Gly Asp
            180                 185                 190

Pro Met Gln Ala Asp Gly Pro Arg Gly Cys Arg Phe Ala Gly Gln Trp
        195                 200                 205

Phe Pro Glu Asn Gln Ser Trp His Pro Ser Val Pro Pro Phe Gly Glu
    210                 215                 220

Met Ser Cys Ile Thr Cys Arg Cys Gly Ala Gly Val Pro His Cys Glu
225                 230                 235                 240

Arg Asp Asp Cys Ser Pro Pro Leu Ser Cys Gly Ser Gly Lys Glu Ser
                245                 250                 255

Arg Cys Cys Ser His Cys Thr Ala Gln Arg Ser Ser Glu Thr Arg Thr
            260                 265                 270

Leu Pro Glu Leu Glu Lys Glu Ala Glu His Ser
        275                 280
```

<210> 8
<211> 955
<212> PRT
<213> Homo sapiens

<400> 8

```
Met Pro Ser Leu Pro Ala Pro Pro Ala Pro Leu Leu Leu Leu Gly Leu
  1               5                  10                  15

Leu Leu Leu Gly Ser Arg Pro Ala Arg Gly Ala Gly Pro Glu Pro Pro
             20                  25                  30

Val Leu Pro Ile Arg Ser Glu Lys Glu Pro Leu Pro Val Arg Gly Ala
         35                  40                  45

Ala Gly Cys Thr Phe Gly Gly Lys Val Tyr Ala Leu Asp Glu Thr Trp
     50                  55                  60

His Pro Asp Leu Gly Glu Pro Phe Gly Val Met Arg Cys Val Leu Cys
 65                  70                  75                  80

Ala Cys Glu Ala Pro Gln Trp Gly Arg Arg Thr Arg Gly Pro Gly Arg
                 85                  90                  95

Val Ser Cys Lys Asn Ile Lys Pro Glu Cys Pro Thr Pro Ala Cys Gly
            100                 105                 110

Gln Pro Arg Gln Leu Pro Gly His Cys Cys Gln Thr Cys Pro Gln Glu
        115                 120                 125
```

```
Arg Ser Ser Ser Glu Arg Gln Pro Ser Gly Leu Ser Phe Glu Tyr Pro
    130                 135                 140

Arg Asp Pro Glu His Arg Ser Tyr Ser Asp Arg Gly Glu Pro Gly Ala
145                 150                 155                 160

Glu Glu Arg Ala Arg Gly Asp Gly His Thr Asp Phe Val Ala Leu Leu
                165                 170                 175

Thr Gly Pro Arg Ser Gln Ala Val Ala Arg Ala Arg Val Ser Leu Leu
            180                 185                 190

Arg Ser Ser Leu Arg Phe Ser Ile Ser Tyr Arg Arg Leu Asp Arg Pro
        195                 200                 205

Thr Arg Ile Arg Phe Ser Asp Ser Asn Gly Ser Val Leu Phe Glu His
    210                 215                 220

Pro Ala Ala Pro Thr Gln Asp Gly Leu Val Cys Gly Val Trp Arg Ala
225                 230                 235                 240

Val Pro Arg Leu Ser Leu Arg Leu Leu Arg Ala Glu Gln Leu His Val
                245                 250                 255

Ala Leu Val Thr Leu Thr His Pro Ser Gly Glu Val Trp Gly Pro Leu
            260                 265                 270

Ile Arg His Arg Ala Leu Ala Ala Glu Thr Phe Ser Ala Ile Leu Thr
        275                 280                 285

Leu Glu Gly Pro Pro Gln Gln Gly Val Gly Gly Ile Thr Leu Leu Thr
    290                 295                 300

Leu Ser Asp Thr Glu Asp Ser Leu His Phe Leu Leu Leu Phe Arg Gly
305                 310                 315                 320

Leu Leu Glu Pro Arg Ser Gly Gly Leu Thr Gln Val Pro Leu Arg Leu
                325                 330                 335

Gln Ile Leu His Gln Gly Gln Leu Leu Arg Glu Leu Gln Ala Asn Val
            340                 345                 350

Ser Ala Gln Glu Pro Gly Phe Ala Glu Val Leu Pro Asn Leu Thr Val
        355                 360                 365

Gln Glu Met Asp Trp Leu Val Leu Gly Glu Leu Gln Met Ala Leu Glu
    370                 375                 380

Trp Ala Gly Arg Pro Gly Leu Arg Ile Ser Gly His Ile Ala Ala Arg
385                 390                 395                 400

Lys Ser Cys Asp Val Leu Gln Ser Val Leu Cys Gly Ala Asp Ala Leu
                405                 410                 415

Ile Pro Val Gln Thr Gly Ala Ala Gly Ser Ala Ser Leu Thr Leu Leu
            420                 425                 430

Gly Asn Gly Ser Leu Ile Tyr Gln Val Gln Val Val Gly Thr Ser Ser
        435                 440                 445

Glu Val Val Ala Met Thr Leu Glu Thr Lys Pro Gln Arg Arg Asp Gln
```

```
    450                 455                 460

Arg Thr Val Leu Cys His Met Ala Gly Leu Gln Pro Gly Gly His Thr
465                 470                 475                 480

Ala Val Gly Ile Cys Pro Gly Leu Gly Ala Arg Gly Ala His Met Leu
                485                 490                 495

Leu Gln Asn Glu Leu Phe Leu Asn Val Gly Thr Lys Asp Phe Pro Asp
            500                 505                 510

Gly Glu Leu Arg Gly His Val Ala Ala Leu Pro Tyr Cys Gly His Ser
        515                 520                 525

Ala Arg His Asp Thr Leu Pro Val Pro Leu Ala Gly Ala Leu Val Leu
    530                 535                 540

Pro Pro Val Lys Ser Gln Ala Ala Gly His Ala Trp Leu Ser Leu Asp
545                 550                 555                 560

Thr His Cys His Leu His Tyr Glu Val Leu Leu Ala Gly Leu Gly Gly
                565                 570                 575

Ser Glu Gln Gly Thr Val Thr Ala His Leu Leu Gly Pro Pro Gly Thr
            580                 585                 590

Pro Gly Pro Arg Arg Leu Leu Lys Gly Phe Tyr Gly Ser Glu Ala Gln
        595                 600                 605

Gly Val Val Lys Asp Leu Glu Pro Glu Leu Leu Arg His Leu Ala Lys
    610                 615                 620

Gly Met Ala Ser Leu Leu Ile Thr Thr Lys Gly Ser Pro Arg Gly Glu
625                 630                 635                 640

Leu Arg Gly Gln Val His Ile Ala Asn Gln Cys Glu Val Gly Gly Leu
                645                 650                 655

Arg Leu Glu Ala Ala Gly Ala Glu Gly Val Arg Ala Leu Gly Ala Pro
            660                 665                 670

Asp Thr Ala Ser Ala Ala Pro Pro Val Val Pro Gly Leu Pro Ala Leu
        675                 680                 685

Ala Pro Ala Lys Pro Gly Gly Pro Gly Arg Pro Arg Asp Pro Asn Thr
    690                 695                 700

Cys Phe Phe Glu Gly Gln Gln Arg Pro His Gly Ala Arg Trp Ala Pro
705                 710                 715                 720

Asn Tyr Asp Pro Leu Cys Ser Leu Cys Thr Cys Gln Arg Arg Thr Val
                725                 730                 735

Ile Cys Asp Pro Val Val Cys Pro Pro Pro Ser Cys Pro His Pro Val
            740                 745                 750

Gln Ala Pro Asp Gln Cys Cys Pro Val Cys Pro Glu Lys Gln Asp Val
        755                 760                 765

Arg Asp Leu Pro Gly Leu Pro Arg Ser Arg Asp Pro Gly Glu Gly Cys
    770                 775                 780
```

```
Tyr Phe Asp Gly Asp Arg Ser Trp Arg Ala Ala Gly Thr Arg Trp His
785                 790                 795                 800

Pro Val Val Pro Pro Phe Gly Leu Ile Lys Cys Ala Val Cys Thr Cys
                805                 810                 815

Lys Gly Gly Thr Gly Glu Val His Cys Glu Lys Val Gln Cys Pro Arg
            820                 825                 830

Leu Ala Cys Ala Gln Pro Val Arg Val Asn Pro Thr Asp Cys Cys Lys
        835                 840                 845

Gln Cys Pro Val Gly Ser Gly Ala His Pro Gln Leu Gly Asp Pro Met
    850                 855                 860

Gln Ala Asp Gly Pro Arg Gly Cys Arg Phe Ala Gly Gln Trp Phe Pro
865                 870                 875                 880

Glu Ser Gln Ser Trp His Pro Ser Val Pro Pro Phe Gly Glu Met Ser
                885                 890                 895

Cys Ile Thr Cys Arg Cys Gly Ala Gly Val Pro His Cys Glu Arg Asp
            900                 905                 910

Asp Cys Ser Leu Pro Leu Ser Cys Gly Ser Gly Lys Glu Ser Arg Cys
        915                 920                 925

Cys Ser Arg Cys Thr Ala His Arg Arg Pro Ala Pro Glu Thr Arg Thr
    930                 935                 940

Asp Pro Glu Leu Glu Lys Glu Ala Glu Gly Ser
945                 950                 955


<210> 9
<211> 452
<212> PRT
<213> Homo sapiens

<400> 9
Met Gly Gly Met Lys Tyr Ile Phe Ser Leu Leu Phe Phe Leu Leu Leu
  1               5                  10                  15

Glu Gly Gly Lys Thr Glu Gln Val Lys His Ser Glu Thr Tyr Cys Met
             20                  25                  30

Phe Gln Asp Lys Lys Tyr Arg Val Gly Glu Arg Trp His Pro Tyr Leu
         35                  40                  45

Glu Pro Tyr Gly Leu Val Tyr Cys Val Asn Cys Ile Cys Ser Glu Asn
     50                  55                  60

Gly Asn Val Leu Cys Ser Arg Val Arg Cys Pro Asn Val His Cys Leu
 65                  70                  75                  80

Ser Pro Val His Ile Pro His Leu Cys Cys Pro Arg Cys Pro Glu Asp
                 85                  90                  95

Ser Leu Pro Pro Val Asn Asn Lys Val Thr Ser Lys Ser Cys Glu Tyr
            100                 105                 110
```

```
Asn Gly Thr Thr Tyr Gln His Gly Glu Leu Phe Val Ala Glu Gly Leu
        115                 120                 125

Phe Gln Asn Arg Gln Pro Asn Gln Cys Thr Gln Cys Ser Cys Ser Glu
    130                 135                 140

Gly Asn Val Tyr Cys Gly Leu Lys Thr Cys Pro Lys Leu Thr Cys Ala
145                 150                 155                 160

Phe Pro Val Ser Val Pro Asp Ser Cys Cys Arg Val Cys Arg Gly Asp
                165                 170                 175

Gly Glu Leu Ser Trp Glu His Ser Asp Gly Asp Ile Phe Arg Gln Pro
            180                 185                 190

Ala Asn Arg Glu Ala Arg His Ser Tyr His Arg Ser His Tyr Asp Pro
        195                 200                 205

Pro Pro Ser Arg Gln Ala Gly Gly Leu Ser Arg Phe Pro Gly Ala Arg
    210                 215                 220

Ser His Arg Gly Ala Leu Met Asp Ser Gln Gln Ala Ser Gly Thr Ile
225                 230                 235                 240

Val Gln Ile Val Ile Asn Asn Lys His Lys His Gly Gln Val Cys Val
                245                 250                 255

Ser Asn Gly Lys Thr Tyr Ser His Gly Glu Ser Trp His Pro Asn Leu
            260                 265                 270

Arg Ala Phe Gly Ile Val Glu Cys Val Leu Cys Thr Cys Asn Val Thr
        275                 280                 285

Lys Gln Glu Cys Lys Lys Ile His Cys Pro Asn Arg Tyr Pro Cys Lys
    290                 295                 300

Tyr Pro Gln Lys Ile Asp Gly Lys Cys Cys Lys Val Cys Pro Gly Lys
305                 310                 315                 320

Lys Ala Lys Glu Glu Leu Pro Gly Gln Ser Phe Asp Asn Lys Gly Tyr
                325                 330                 335

Phe Cys Gly Glu Glu Thr Met Pro Val Tyr Glu Ser Val Phe Met Glu
            340                 345                 350

Asp Gly Glu Thr Thr Arg Lys Ile Ala Leu Glu Thr Glu Arg Pro Pro
        355                 360                 365

Gln Val Glu Val His Val Trp Thr Ile Arg Lys Gly Ile Leu Gln His
    370                 375                 380

Phe His Ile Glu Lys Ile Ser Lys Arg Met Phe Glu Glu Leu Pro His
385                 390                 395                 400

Phe Lys Leu Val Thr Arg Thr Thr Leu Ser Gln Trp Lys Ile Phe Thr
                405                 410                 415

Glu Gly Glu Ala Gln Ile Ser Gln Met Cys Ser Ser Arg Val Cys Arg
            420                 425                 430
```

Thr Glu Leu Glu Asp Leu Val Lys Val Leu Tyr Leu Glu Arg Ser Glu
435 440 445

Lys Gly His Cys
450

<210> 10
<211> 11
<212> PRT
<213> Human immunodeficiency virus type 1

<400> 10
Tyr Gly Arg Lys Lys Arg Arg Gln Arg Arg Arg
1 5 10

<210> 11
<211> 15
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: internalizing domain derived from HIV tat protein

<400> 11
Gly Gly Gly Gly Tyr Gly Arg Lys Lys Arg Arg Gln Arg Arg Arg
1 5 10 15

<210> 12
<211> 23
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: oligonucleotide; primer 1605-21

<400> 12
aatccgatgc ccacgttgca gta 23

<210> 13
<211> 26
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: oligonucleotide; PCR primer 1239-08

<400> 13
aaaatcttag accgacgact gtgttt 26

<210> 14
<211> 25
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:
oligonucleotide; PCR primer

<400> 14
cgtaaaagat cctgcgctag atgcg 25

<210> 15
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:
oligonucleotide; PCR primer

<400> 15
tcctctcatc ctcaccttag 20

<210> 16
<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:
oligonucleotide; PCR primer

<400> 16
ggagaaagtg agataaggac ac 22

<210> 17
<211> 40
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:
oligonucleotide; PCR primer 2360-40

<400> 17
gctatctaga gccaccatgg ttcccggggt gaggatcatc 40

<210> 18
<211> 36
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:
oligonucleotide; PCR primer 2360-41

<400> 18
gctagtcgac ctataatgtc ttggtcactt tgtctg 36

<210> 19
<211> 42
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:
oligonucleotide; PCR primer

<400> 19
gctagcggcc gcgccaccat ggttcccggg gtgaggatca tc 42

<210> 20
<211> 36
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:
oligonucleotide; PCR primer

<400> 20
gctagtcgac taatgtcttg gtcactttgt ctgggc 36

<210> 21
<211> 17
<212> PRT
<213> Mus musculus

<400> 21
Cys Pro Glu Asp Glu Ala Glu Asp Asp His Ser Glu Val Ile Ser Thr
1 5 10 15

Arg

# INTERNATIONAL SEARCH REPORT

International Application No
PCT/US 01/06891

A. CLASSIFICATION OF SUBJECT MATTER
IPC 7 C12N15/12 C07K14/47 C07K14/705 C12N15/62 C07K16/18
C07K16/28 G01N33/53 A61K38/17 C12Q1/68

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 7 C12N C07K G01N A61K C12Q

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

| Category ° | Citation of document, with indication, where appropriate, of the relevant passages | Relevant to claim No. |
|---|---|---|
| X | WO 00 09551 A (GENETICS INST)<br>24 February 2000 (2000-02-24)<br><br>* see passages relating to clone dw665_4 *<br>the whole document<br>--- | 1-12,<br>14-18,<br>20-29,<br>31-33,<br>35-38,<br>40-56 |
| X | WO 99 54353 A (SCHMITT ARMIN ;SPECHT THOMAS (DE); DAHL EDGAR (DE); HINZMANN BERND) 28 October 1999 (1999-10-28)<br><br>* see seqID's 4, 19, 72 and 113 *<br>---<br>-/-- | 1-12,<br>14-32,<br>34,35,<br>37,38,<br>40,43,<br>46,56 |

[X] Further documents are listed in the continuation of box C.

[X] Patent family members are listed in annex.

° Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

| Date of the actual completion of the international search | Date of mailing of the international search report |
|---|---|
| 7 August 2001 | 21/08/2001 |

| Name and mailing address of the ISA | Authorized officer |
|---|---|
| European Patent Office, P.B. 5818 Patentlaan 2<br>NL - 2280 HV Rijswijk<br>Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,<br>Fax: (+31-70) 340-3016 | Smalt, R |

Form PCT/ISA/210 (second sheet) (July 1992)

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

| Category * | Citation of document, with indication,where appropriate, of the relevant passages | Relevant to claim No. |
|---|---|---|
| P,X | WO 00 12708 A (BAKER KEVIN ;GENENTECH INC (US); GODDARD AUDREY (US); GURNEY AUSTI) 9 March 2000 (2000-03-09) * see PRO1557 (43rd cloned seq.) * | 1-12, 14-40, 43-56 |
| E | WO 01 29084 A (LEXICON GENETICS INC) 26 April 2001 (2001-04-26) the whole document | 1-40, 43-55 |
| E | WO 01 34796 A (SAVITZKY KINNERET ;TOPOROIK AMIR (IL); BITON SHARON (IL); COMPUGEN) 17 May 2001 (2001-05-17) the whole document | 1-12, 14-40, 43-56 |
| T | NAKAYAMA NAOKI ET AL: "A novel chordin-like protein inhibitor for bone morphogenetic proteins expressed preferentially in mesenchymal cell lineages." DEVELOPMENTAL BIOLOGY, vol. 232, no. 2, 15 April 2001 (2001-04-15), pages 372-387, XP002174073 ISSN: 0012-1606 the whole document | |

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

Continuation of Box I.2

Present claims 18, 19, 31, 32 and 34 relate to a product defined by reference to a desirable characteristic or property, namely its affinity for the polypeptide(s) of the invention.
The claims cover all products having this characteristic or property, whereas the application provides support within the meaning of Article 6 PCT and/or disclosure within the meaning of Article 5 PCT for only a very limited number of such compounds. In the present case, the claims so lack support, and the application so lacks disclosure, that a meaningful search over the whole of the claimed scope is impossible. Independent of the above reasoning, the claims also lack clarity (Article 6 PCT). An attempt is made to define the product by reference to a result to be achieved. Again, this lack of clarity in the present case is such as to render a meaningful search over the whole of the claimed scope impossible. Consequently, the search has been carried out for those parts of the claims which appear to be clear, supported and disclosed, namely those parts relating to antibodies directed at the polypeptide(s) of the invention, as they can to some extent be characterized by the protein to which they bind.

The applicant's attention is drawn to the fact that claims, or parts of claims, relating to inventions in respect of which no international search report has been established need not be the subject of an international preliminary examination (Rule 66.1(e) PCT). The applicant is advised that the EPO policy when acting as an International Preliminary Examining Authority is normally not to carry out a preliminary examination on matter which has not been searched. This is the case irrespective of whether or not the claims are amended following receipt of the search report or during any Chapter II procedure.

# INTERNATIONAL SEARCH REPORT

Interna[illegible]l Application No
PCT/US 01/06891

| Patent document cited in search report | Publication date | Patent family member(s) | Publication date |
|---|---|---|---|
| WO 0009551 A | 24-02-2000 | AU 4071199 A | 23-11-1999 |
| | | AU 5475199 A | 06-03-2000 |
| | | EP 1077991 A | 28-02-2001 |
| | | EP 1112285 A | 04-07-2001 |
| | | WO 9957132 A | 11-11-1999 |
| WO 9954353 A | 28-10-1999 | DE 19817946 A | 21-10-1999 |
| | | EP 1071777 A | 31-01-2001 |
| WO 0012708 A | 09-03-2000 | AU 5590899 A | 21-03-2000 |
| | | AU 6041399 A | 10-04-2000 |
| | | EP 1115863 A | 18-07-2001 |
| | | WO 0017353 A | 30-03-2000 |
| WO 0129084 A | 26-04-2001 | NONE | |
| WO 0134796 A | 17-05-2001 | NONE | |

# INTERNATIONAL SEARCH REPORT

International Application No
PCT/US 01/06891

A. CLASSIFICATION OF SUBJECT MATTER
IPC 7 C12N15/12 C07K14/47 C07K14/705 C12N15/62 C07K16/18
C07K16/28 G01N33/53 A61K38/17 C12Q1/68

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 7 C12N C07K G01N A61K C12Q

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

| Category ° | Citation of document, with indication, where appropriate, of the relevant passages | Relevant to claim No. |
|---|---|---|
| X | WO 00 09551 A (GENETICS INST) 24 February 2000 (2000-02-24)<br><br>* see passages relating to clone dw665_4 *<br>the whole document | 1-12, 14-18, 20-29, 31-33, 35-38, 40-56 |
| X | WO 99 54353 A (SCHMITT ARMIN ;SPECHT THOMAS (DE); DAHL EDGAR (DE); HINZMANN BERND) 28 October 1999 (1999-10-28)<br><br>* see seqID's 4, 19, 72 and 113 *<br><br>-/-- | 1-12, 14-32, 34,35, 37,38, 40,43, 46,56 |

[X] Further documents are listed in the continuation of box C.

[X] Patent family members are listed in annex.

° Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

| Date of the actual completion of the international search | Date of mailing of the international search report |
|---|---|
| 7 August 2001 | 21/08/2001 |

| Name and mailing address of the ISA | Authorized officer |
|---|---|
| European Patent Office, P.B. 5818 Patentlaan 2<br>NL - 2280 HV Rijswijk<br>Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,<br>Fax: (+31-70) 340-3016 | Smalt, R |

Form PCT/ISA/210 (second sheet) (July 1992)

# INTERNATIONAL SEARCH REPORT

International Application No
PCT/US 01/06891

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

| Category ° | Citation of document, with indication, where appropriate, of the relevant passages | Relevant to claim No. |
|---|---|---|
| P,X | WO 00 12708 A (BAKER KEVIN ;GENENTECH INC (US); GODDARD AUDREY (US); GURNEY AUSTI) 9 March 2000 (2000-03-09) * see PRO1557 (43rd cloned seq.) * | 1-12, 14-40, 43-56 |
| E | WO 01 29084 A (LEXICON GENETICS INC) 26 April 2001 (2001-04-26) the whole document | 1-40, 43-55 |
| E | WO 01 34796 A (SAVITZKY KINNERET ;TOPOROIK AMIR (IL); BITON SHARON (IL); COMPUGEN) 17 May 2001 (2001-05-17) the whole document | 1-12, 14-40, 43-56 |
| T | NAKAYAMA NAOKI ET AL: "A novel chordin-like protein inhibitor for bone morphogenetic proteins expressed preferentially in mesenchymal cell lineages." DEVELOPMENTAL BIOLOGY, vol. 232, no. 2, 15 April 2001 (2001-04-15), pages 372-387, XP002174073 ISSN: 0012-1606 the whole document | |

Form PCT/ISA/210 (continuation of second sheet) (July 1992)

# INTERNATIONAL SEARCH REPORT

International application No.
PCT/US 01/06891

**Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)**

This International Search Report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. [X] Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

Although claims 33, 48, and 53 are directed to a method of treatment of the human/animal body, the search has been carried out and based on the alleged effects of the compound/composition.

2. [X] Claims Nos.:
because they relate to parts of the International Application that do not comply with the prescribed requirements to such an extent that no meaningful International Search can be carried out, specifically:

see FURTHER INFORMATION sheet PCT/ISA/210

3. [ ] Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

**Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)**

This International Searching Authority found multiple inventions in this international application, as follows:

1. [ ] As all required additional search fees were timely paid by the applicant, this International Search Report covers all searchable claims.

2. [ ] As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.

3. [ ] As only some of the required additional search fees were timely paid by the applicant, this International Search Report covers only those claims for which fees were paid, specifically claims Nos.:

4. [ ] No required additional search fees were timely paid by the applicant. Consequently, this International Search Report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

[ ] The additional search fees were accompanied by the applicant's protest.

[ ] No protest accompanied the payment of additional search fees.

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

Continuation of Box I.2

Present claims 18, 19, 31, 32 and 34 relate to a product defined by reference to a desirable characteristic or property, namely its affinity for the polypeptide(s) of the invention.
The claims cover all products having this characteristic or property, whereas the application provides support within the meaning of Article 6 PCT and/or disclosure within the meaning of Article 5 PCT for only a very limited number of such compounds. In the present case, the claims so lack support, and the application so lacks disclosure, that a meaningful search over the whole of the claimed scope is impossible. Independent of the above reasoning, the claims also lack clarity (Article 6 PCT). An attempt is made to define the product by reference to a result to be achieved. Again, this lack of clarity in the present case is such as to render a meaningful search over the whole of the claimed scope impossible. Consequently, the search has been carried out for those parts of the claims which appear to be clear, supported and disclosed, namely those parts relating to antibodies directed at the polypeptide(s) of the invention, as they can to some extent be characterized by the protein to which they bind.

The applicant's attention is drawn to the fact that claims, or parts of claims, relating to inventions in respect of which no international search report has been established need not be the subject of an international preliminary examination (Rule 66.1(e) PCT). The applicant is advised that the EPO policy when acting as an International Preliminary Examining Authority is normally not to carry out a preliminary examination on matter which has not been searched. This is the case irrespective of whether or not the claims are amended following receipt of the search report or during any Chapter II procedure.

# INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No
PCT/US 01/06891

| Patent document cited in search report | Publication date | Patent family member(s) | Publication date |
|---|---|---|---|
| WO 0009551 A | 24-02-2000 | AU 4071199 A | 23-11-1999 |
| | | AU 5475199 A | 06-03-2000 |
| | | EP 1077991 A | 28-02-2001 |
| | | EP 1112285 A | 04-07-2001 |
| | | WO 9957132 A | 11-11-1999 |
| WO 9954353 A | 28-10-1999 | DE 19817946 A | 21-10-1999 |
| | | EP 1071777 A | 31-01-2001 |
| WO 0012708 A | 09-03-2000 | AU 5590899 A | 21-03-2000 |
| | | AU 6041399 A | 10-04-2000 |
| | | EP 1115863 A | 18-07-2001 |
| | | WO 0017353 A | 30-03-2000 |
| WO 0129084 A | 26-04-2001 | NONE | |
| WO 0134796 A | 17-05-2001 | NONE | |